BELLMAN FUNCTIONS TECHNIQUE IN HARMONIC ANALYSIS

VASILY VASYUNIN AND ALEXANDER VOLBERG

1. Introduction.

2. Bellman functions and Martinagales

We will be working with stochastic integrals. But we do not need subtle facts of the theory of stochastic integrals. We need just to use some formal rules of working with them. Basically the only things we will use is Itô's isometry and Itô's formula. The standard reference is [StInt], [Ve].

We start with two test functions f, g on $\mathbb{R}^d$. Let U_f, U_g mean the heat extensions of them into $\mathbb{R}^d \times \mathbb{R}_+$, i.e. $U_f(x,t) := \mathbb{E} f(x + W_t)$ (and similarly for U_g), where $(W_t)_{t \geq 0}$ is a standard d-dimensional Brownian motion starting at the origin. Hence, $\frac{\partial U_f}{\partial t} = \frac{1}{2} \Delta U_f$ and $U_f(x, 0) = f(x)$ for $x \in \mathbb{R}^d$.

Let us fix $T > 0$. In what follows we will be using a simple key remark:

$$F_t := U_f(W_t, T - t) \text{ is a martingale}\,. \tag{2.1}$$

Also

$$F_t - F_0 = \int_0^t \nabla_x U_f(W_s, T - s) \cdot dW_s\,. \tag{2.2}$$

Let us abuse the language and call such martingales "heat martingales", they are always obtained by the heat extension of a test function.

Next we remark that *these are not all possible interesting martingales with which we will have to work*. The fact is that we will sometimes need *the martingale transforms of heat martingales*. To explain what is it, let us fix a $d \times d$ matrix $A = (a_{kl})$ and consider a new martingale

$$(A \star F)_t = \int_0^t A \nabla_x U_f(W_s, T - s) \cdot dW_s\,. \tag{2.3}$$

For convinience denote $Z_t := (W_t, T - t)$ (sometimes it is called *space-time Brownian motion*, it was used first by Varopoulos [Varo1] and later by [BaJa1], [BaJa2].

Consider the operator

$$(T_A f)(x) := \lim_{T\to\infty} \mathbb{E}(A \star F | Z_T = (x,0)) . \tag{2.4}$$

The first task, to which we want to apply Bellman function technique, is the estimate of certain polynomials of Riesz transforms. Let us remind that the Riesz transform R_j is given by

$$R_j f = \mathcal{F}^{-1} i\xi_j \mathcal{F} f\,,\ j = 1, \cdots, d\,.$$

Here is a simple but important theorem (see [GMSS]).

Theorem 1. *$T_A = -\sum_{k,l=1}^d a_{kl} R_k\, R_l$, where R_i are Riesz transforms on $\mathbb{R}^d$.*

Proof. By definition

$$T_A f(x) = \lim_{T\to\infty} \mathbb{E}_{\{W_T = x\}}(\int_0^T A\nabla U_f(W_t, T-t)\cdot dW_t)\,.$$

Let us denote (temporarily) by $(T_A\, f, g)$ the bilinear form $\int T_A\, f(x)\cdot g(x)dx$. Then

$$(T_A\, f, g) = \int dx \lim_{T\to\infty} \mathbb{E}_{\{W_T = x\}}(\int_0^T A\nabla U_f(W_t, T-t)\cdot dW_t) g(x) =$$

$$\int dx \lim_{T\to\infty} \mathbb{E}_{\{W_T = x\}}(g(W_T)\int_0^T A\nabla U_f(W_t, T-t)\cdot dW_t) =$$

$$\lim_{T\to\infty}\int d\mu_T(x)\,(2\pi T)^{\frac{d}{2}}\mathbb{E}_{\{W_T = x\}}(g(W_T)\int_0^T A\nabla U_f(W_t, T-t)\cdot dW_t) =$$

(the dominant convergence theorem as f, g are bounded with compact support, $d\mu_T$ is the distribution measure for W_T, that is $\frac{1}{(2\pi T)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2T}}\,dx$)

$$= \lim_{T\to\infty}\int d\mu_T(x)(2\pi T)^{\frac{d}{2}}\,\mathbb{E}_{\{W_T = x\}}(\int_0^T \nabla U_g(W_t, T-t)\cdot dW_t \int_0^T A\nabla U_f(W_t, T-t)\cdot dW_t)$$

as by (2.2)

$$g(W_T) - U_g(0,T) = \int_0^T \nabla U_g(W_t, T-t)\cdot dW_t\,,$$

and $\lim_{T\to\infty} U_g(0,T) = 0$.

Then $\int d\mu_T(x)\,\mathbb{E}_{\{W_T = x\}}(\cdot) =: \mathbb{E}(\cdot)$. This and the previous chain of equalities give us

Therefore,

$$(T_A f, g) = \lim_{T\to\infty}(2\pi T)^{\frac{d}{2}}\mathbb{E}(\int_0^T \nabla U_g(W_t, T-t)\cdot dW_t \int_0^T A\nabla U_f(W_t, T-t)\cdot dW_t)\,.$$

By Itô's isometry (this is a simple fact, a standard reference is [StInt], [Ve])

$$(T_A f, g) = \lim_{T\to\infty} (2\pi T)^{\frac{d}{2}} \int_0^T \mathbb{E}\langle A\nabla U_f(W_t, T-t), \nabla U_g(W_t, T-t)\rangle \, dt \,.$$

Here $\langle\cdot,\cdot\rangle$ is again a linear form (now on vectors in $\mathbb{C}^d$). Denote

$$I_T := (2\pi T)^{\frac{d}{2}} \mathbb{E}\langle A\nabla U_f(W_t, T-t), \nabla U_g(W_t, T-t)\rangle \, dt \,.$$

Then

$$I_T = (2\pi T)^{\frac{d}{2}} \int_0^T \int_{\mathbb{R}^d} \langle A\nabla_x U_f(x,t), \nabla_x U_g(x,t)\rangle d\mu_{t,T}(x)\, dt \,,$$

where $d\mu_{t,T}$ is the law of W_{T-t}. Now we can easily see that

$$\lim_{T\to\infty} I_T = \int_{\mathbb{R}^d_+} \langle A\nabla_x U_f(x,t), \nabla_x U_g(x,t)\rangle dx\, dt \,.$$

But the Fourier taransform $(\mathcal{F}U_f)(\xi,t) = (\mathcal{F}f)(\xi)e^{-t|\xi|^2/2}$. So, by Parceval's formula

$$\lim_{T\to\infty} I_T = \int_{\mathbb{R}^d_+} (Ai\xi, i\xi)\mathcal{F}f(\xi)\overline{\mathcal{F}\bar{g}(\xi)} e^{-t|\xi|^2} d\xi\, dt = \int_{\mathbb{R}^d} \frac{(A\xi,\xi)}{|\xi|^2} \mathcal{F}f(\xi)\overline{\mathcal{F}\bar{g}(\xi)}\, d\xi \,.$$

Taking into consideration that *bilinear* form $(R_k\, R_l f, g) = -\int_{\mathbb{R}^d} \frac{\xi_k\xi_l}{|\xi|^2}\mathcal{F}f(\xi)\overline{\mathcal{F}\bar{g}(\xi)}\, d\xi$ just by definition of the Riesz transforms, we finish the proof of the theorem by inspecting two last equalities. □

2.1. **Examples.** Operator $T_1 := R_1^2 - R_2^2$ corresponds to matrix $A_1 = \begin{bmatrix} -1, & 0 \\ 0, & 1 \end{bmatrix}$.

Operator $T_2 = -2iR_1R_2$ corresponds to matrix $A_2 = \begin{bmatrix} 0, & i \\ i, & 0 \end{bmatrix}$.

Ahlfors-Beurling operator $T = R_1^2 - R_2^2 - 2iR_1R_2$ corresponds to matrix $A = \begin{bmatrix} -1, & i \\ i, & 1 \end{bmatrix}$.

Matrix norms are correspondingly

$$\|A_1\| = \|A_2\| = 1\,, \|A\| = 2\,. \tag{2.5}$$

2.2. **Two useful formulae.** While proving Theorem 1 we deduced two useful formulae. The first one is

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar{g}(x)\, dx =$$
$$\int_{\mathbb{R}^d_+} \sum_{k,l=1}^{d} a_{kl}\partial_{x_k} U_f(x,t)\partial_{x_l} U_{\bar{g}}(x,t)\, dx\, dt = \int_{\mathbb{R}^d_+} \sum_{k,l=1}^{d} A\nabla U_f \cdot \nabla U_{\bar{g}}\, dx\, dt\,, \tag{2.6}$$

where $\cdot$ stands for the scalar product in $\mathbb{C}^d$.

The second formula is

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar{g}(x)\, dx =$$
$$\lim_{T\to\infty} \mathcal{E}\{(A\star F)(T)\,\bar{G}(T)\} = \lim_{T\to\infty} (2\pi T)^{\frac{d}{2}}\mathbb{E}\{(A\star F)(T)\,\bar{G}(T)\}\,, \tag{2.7}$$

where $\mathcal{E}$ is the expectation not with respect to probability measure $d\mu_T(x)dP^T_x$, but with respect to *pseudo-probability* measure $dxdP^T_x$ (dP^T_x is the conditional probability measure that on $W_T = x$) and where

$$G(T) = \int_0^T \nabla U_g(W_s, T-s)\cdot dW_s\,. \tag{2.8}$$

Just separating real and imaginary parts one gets

$$G(T) = X_1(T) + iX_2(T)\,, \tag{2.9}$$

and

$$X_1(T) = \int_0^T \overrightarrow{H_1}(s)\cdot dW_s\,,$$
$$X_2(T) = \int_0^T \overrightarrow{H_2}(s)\cdot dW_s\,,$$

where X_1, X_2 are *real-valued* processes, and $\overrightarrow{H_1}(s), \overrightarrow{H_2}(s)$ are $\mathbb{R}^d$–valued martingales adapted to the filtration of d-dimensional Brownian motion W_s. We can easily write components of $\overrightarrow{H_1}(s), \overrightarrow{H_2}(s)$:

$$H^i_1(s) = \partial_{x_i} U_{\Re g}(W_s, T-s)\,,\;\; H^i_2(s) = \partial_{x_i} U_{\Im g}(W_s, T-s)\,, i = 1, ..., d\,.$$

Similarly

$$(A\star F)(T) = Y_1(T) + iY_2(T)\,, \tag{2.10}$$
$$Y_1(T) = \int_0^T \overrightarrow{K_1}(s)\cdot dW_s\,,\, Y_2(T) = \int_0^T \overrightarrow{K_2}(s)\cdot dW_s\,,$$

where Y_1, Y_2 are *real-valued* processes, and $\overrightarrow{K_1}(s), \overrightarrow{K_2}(s)$ are $\mathbb{R}^d$–valued martingales adapted to the filtration of d-dimensional Brownian motion W_s. We can easily write components of $\overrightarrow{K_1}(s), \overrightarrow{K_2}(s)$:

$$K_1^i(s) = \Re \sum_{k=1}^{d} a_{ik}\partial_{x_k}\ U_f(W_s, T-s)\,,\ K_2^i(s) = \Im \sum_{k=1}^{d} a_{ik}\partial_{x_k} U_f(W_s, T-s)\,.$$

Notice that for matrix A (and $d=2$)

$$\overrightarrow{K_1} = \begin{bmatrix} (\Re f_x - \Im f_y)(W_s, T-s) \\ (-\Re f_y - \Im f_x)(...) \end{bmatrix}\,,\ \overrightarrow{K_2} = \begin{bmatrix} (\Re f_y + \Im f_x)(W_s, T-s) \\ (\Re f_x - \Im f_y)(...) \end{bmatrix}\,. \tag{2.11}$$

After these notations (2.7) can be rewritten as follows:

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar{g}(x)\, dx = \lim_{T\to\infty} \mathcal{E} \int_0^T (\overrightarrow{K_1}(s)\cdot\overrightarrow{H_1}(s) + \overrightarrow{K_2}(s)\cdot\overrightarrow{H_2}(s))\, ds\,. \tag{2.12}$$

Here $\cdot$ means the scalar product of the corresponding d–vectors.

2.3. **Local ortogonality.** The processes

$$\langle X_i, Y_j\rangle(t) := \int_0^t \overrightarrow{H_i}\cdot\overrightarrow{K_j} ds\,,\ i,j = 1,2$$

$$\langle X_i, X_j\rangle(t) := \int_0^t \overrightarrow{H_i}\cdot\overrightarrow{H_j} ds\,,\ i,j = 1,2$$

$$\langle Y_i, Y_j\rangle(t) := \int_0^t \overrightarrow{K_i}\cdot\overrightarrow{K_j} ds\,,\ i,j = 1,2$$

are called the covariance processes. We can denote

$$d\langle X_i, Y_j\rangle(t) := \overrightarrow{H_i}(t)\cdot\overrightarrow{K_j}(t)\,,\ i,j = 1,2$$

$$d\langle X_i, X_j\rangle(t) := \overrightarrow{H_i}(t)\cdot\overrightarrow{H_j}(t)\,,\ i,j = 1,2$$

$$d\langle Y_i, Y_j\rangle(t) := \overrightarrow{K_i}(t)\cdot\overrightarrow{K_j}(t)\,,\ i,j = 1,2$$

So (2.12) becomes

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar{g}(x)\, dx = \lim_{T\to\infty} \mathcal{E} \int_0^T (d\langle X_1, Y_1\rangle(t) + d\langle X_2, Y_2\rangle(t))\, dt\,. \tag{2.13}$$

Important is

Lemma 2. *Let* $A = \begin{bmatrix} -1, & i \\ i, & 1 \end{bmatrix}$. *Then*

$$d\langle Y_1, Y_2\rangle(t) = 0\,.$$

Or

$$\overrightarrow{K}_1(t)\cdot\overrightarrow{K}_2(t) = 0\,.$$

Proof. It is obvious from (2.11). □

3. Monotone functionals and Bellman function.

So far we introduced martingales X_1, X_2, Y_1, Y_2 as in (2.8), (2.9), (2.10). This was done with the aim at estimating $(T_A\, f, g)$. We will need also

$$P(t) := U_{|f|^p}(W_t, T-t)\,, \tag{3.1}$$

$$Q(t) := U_{|g|^q}(W_t, T-t)\,, \tag{3.2}$$

We repeat that these are martingales.

To estimate (see (2.13))

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar{g}(x)\,dx = \lim_{T\to\infty} \mathcal{E}\int_0^T (d\langle X_1, Y_1\rangle(t) + d\langle X_2, Y_2\rangle(t))\,dt$$

we will use "a monotone functional" approach.

We want to invent a function B of 6 real variables such that the functional

$$\Phi(t) = \mathcal{E}\{B(P(t), X_1(t), X_2(t), Q(t), Y_1(t), Y_2(t))\}$$

were monotone in t. Having in mind (2.13) and having in mind our wish to prove (with the best possible $C(p)$)

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar{g}(x)\,dx \le C(p)\|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^q(\mathbb{R}^d)}\,.$$

we postulate the following desired properties of B.

Desired properties of B.

1)

$$-\Phi'(t) \ge \mathcal{E}(|d\langle X_1, Y_1\rangle(t)| + |d\langle X_2, Y_2\rangle(t)|)\,.$$

or stronger

1')

$$-\Phi'(t) \ge \mathcal{E}(|\overrightarrow{K_1}(t)||\overrightarrow{H_1}(t)| + |\overrightarrow{K_2}(t)||\overrightarrow{H_2}(t)|)\,.$$

2)

$$0 \le B(P, X_1, X_2, Q, Y_1, Y_2) \le C(p) P^{1/p} Q^{1/q}\,.$$

Notice that from 1), and 2) it follows that

$$C(p)\|A\|\|f\|_p\|g\|_q \ge \mathcal{E}\Phi(0) \ge -\mathcal{E}\int_0^T \Phi'(t)\,dt + \mathcal{E}\Phi(T)\,.$$

And so

$$C(p)\|A\|\|f\|_p\|g\|_q \ge \mathcal{E}\int_0^T (d\langle X_1, Y_1\rangle(t) + d\langle X_2, Y_2\rangle(t))\,dt\,.$$

Taking T to infinity and comparing with (2.13) we get the estimate of our polynomial of Riesz transforms T_A:

$$C(p)\|A\|\|f\|_p\|g\|_q \ge |(T_A\, f, g)|\,. \tag{3.3}$$

Property 2) is in terms of B. Property 1), on the other hand, is in terms of expectation rather than in terms of B itself. Property 2) is in terms of B.Can we find a condition on B that garantees property 1)? This would be nice because then we are able to solve for B! Actually we can indeed find condition on B ensuring 1). It is done with the help of Itô's formula.

3.1. **Itô's formula.** We need to change the notations temporarily. This is almost inevitable as $B(P, X_1, X_2, Q, Y_1, Y_2)$ becomes awkward, all variables have different names. In this section we will call them $a = (a_1, a_2, ..., a_6)$.

Here is a Itô's formula (a standard reference is [StInt], [Ve]):

$$dB(a(t)) = \langle \nabla B(a(t)), d\,a(t)\rangle + \frac{1}{2}\sum_{i,j=1}^{6} \frac{\partial^2 B(a(t))}{\partial a_i \partial a_j} d\langle a_i, a_j\rangle(t)\,.$$

Here $\langle\cdot,\cdot\rangle$ is the scalar product in $\mathbb{R}^6$. The beauty of this formula is in particular in the fact that the first term is a martingale difference process and so its expectation iz zero.

Therefore,

$$-\Phi'(t) = -\mathcal{E} dB(a(t)) = -\frac{1}{2}\mathcal{E}\sum_{i,j=1}^{6} \frac{\partial^2 B(a(t))}{\partial a_i \partial a_j} d\langle a_i, a_j\rangle(t)\,. \tag{3.4}$$

Let us now use the fact that all our maringales $a_i(t)$ have the representation

$$a_i(t) - a_i(0) = \int_0^t \overrightarrow{\alpha_i}(s)\cdot dW_s\,,$$

where $\alpha_i(s) := (\alpha_i^1, ..., \alpha_i^d)$.

We will also use the notations

$$\alpha^k := (\alpha_1^k, ..., \alpha_6^k)\,,\ k = 1, .., d\,.$$

In the previous notations $\overrightarrow{K_1} = \overrightarrow{\alpha_5}$, $\overrightarrow{K_2} = \overrightarrow{\alpha_6}$, $\overrightarrow{H_1} = \overrightarrow{\alpha_2}$, $\overrightarrow{H_2} = \overrightarrow{\alpha_3}$.

In particular, local orthogonality (2) that one gets for particular case $d = 2, A = \begin{bmatrix} -1, & i \\ i, & 1 \end{bmatrix}$ implies

$$\overrightarrow{\alpha_5}(s) \cdot \overrightarrow{\alpha_6}(s) = 0\,. \tag{3.5}$$

Now expectation of Itô's formula can be rewritten

$$-\Phi'(t) = -\mathcal{E}dB(a(t)) = \frac{1}{2}\mathcal{E}\sum_{k=1}^{d}\sum_{i,j=1}^{6}\frac{\partial^2 B(a(t))}{\partial a_i \partial a_j}\alpha_i^k\alpha_j^k(t) = \frac{1}{2}\sum_{k=1}^{d}\langle -\frac{d^2B}{da^2}\alpha^k, \alpha^k\rangle\,. \tag{3.6}$$

Here $\langle\cdot,\cdot\rangle$ is the scalar product in $\mathbb{R}^6$.

Differential property of B, concavity of B.

From (3.6) it is easy to have a condition on B ensuring property 1) above.

Suppose that B satisfies two conditions:

I) For an arbitrary $\alpha \in \mathbb{R}^6$ we want to have

$$\langle -\frac{d^2B}{da^2}\alpha, \alpha\rangle \geq 2((\alpha_2)^2 + (\alpha_3)^2)^{1/2}((\alpha_5)^2 + (\alpha_6)^2)^{1/2} \geq 2(|\alpha_2||\alpha_5| + |\alpha_3||\alpha_6|)\,.$$

We rewrite property 2) in a more precise form: II)

$$B(a) \leq (p^* - 1)a_1^{1/p}a_4^{1/p'} textwherep^* = max(p, p'$$

Theorem 3. *Function B satisfying properties I) and II) exists in the domain $\Omega = \{...\}$. Factor 2 in the right hand side cannot be increased, constant $p^* - 1$ cannot be decreased.*

This theorem will be proved in Section 7.

It has powerful consequences. First let us mix the notations:

$$d\langle X_1, Y_1\rangle = \sum_{k=1}^{d}\alpha_2^k\alpha_5^k\,,\ \ d\langle X_2, Y_2\rangle = \sum_{k=1}^{d}\alpha_3^k\alpha_6^k\,.$$

Using this, summing up property I) for $\alpha := \alpha^k, k = 1, .., d$ and applying (3.6) we get

$$-\Phi'(t) \geq \mathcal{E}(d\langle X_1, Y_1\rangle + d\langle X_2, Y_2\rangle) \tag{3.7}$$

and we get (3.3), with $A := A_1 = \begin{bmatrix} -1, & 0 \\ 0, & 1 \end{bmatrix}$, $A := A_2 = \begin{bmatrix} 0, & i \\ i, & 0 \end{bmatrix}$, $A = \begin{bmatrix} -1, & i \\ i, & 1 \end{bmatrix}$.

Consider T_A in $\mathbb{R}^d, d \geq 2$, and A being three matrices introduced above. Then we have $T_{A_1} = R_1^2 - R_2^2,\ T_{A_2} = -2iR_1R_2,\ T_A = T_{A_1} + iT_{A_2}$.

Of course $\mathcal{B} := T_A$ is the Ahlfors-Beurling operator.

We will prove soon the following result.

Theorem 4.

$$\|T_{A_1}\| = \|T_{A_2}\| \leq p^* - 1\,,$$

$$\|T_A\| \leq 2(p^* - 1)\,.$$

A natural question arises: how sharp is theorem 4. In Section 11 we will prove that

$$\|T_{A_1}\| = \|T_{A_2}\| \geq p^* - 1\,,\ \|T_A\| \geq p^* - 1\,.$$

In particular we get an exact norm of $R_1^2 - R_2^2$ and of R_1R_2:

Theorem 5.

$$\|R_1^2 - R_2^2\| = p^* - 1\,,$$

$$\|R_1R_2\| = \frac{p^* - 1}{2}\,.$$

3.2. **Estimate from above for $\|T_A\|$. Calculation of the norm of the quantum combination of $R_1^2 - R_2^2$ and $2R_1R_2$.** Write $\mathcal{B}$ as $\Re\mathcal{B} + i\Im\mathcal{B}$ by splitting its kernel to real and imaginary parts. Of course $\Re\mathcal{B} = R_1^2 - R_2^2$, $\Im\mathcal{B} = 2R_1R_2$. Theorem 5 claims that any convex combination of $\Re\mathcal{B} = R_1^2 - R_2^2$ and $\Im\mathcal{B} = 2R_1R_2$ has norm at most $p^* - 1$. However, we wish to prove that any "quantum convex" combination has norm exactly $p^* - 1$.

Given $\theta \in [0, 2\pi)$ let us introduce the matrix $A^\theta = \begin{bmatrix} \cos\theta, & \sin\theta \\ \sin\theta, & -\cos\theta \end{bmatrix}$. This is the reflection with respect to $0x$ axis followed by the rotation. We saw already that this matrix A^θ corresponds to operator

$$T_{A^\theta} = (R_1^2 - R_2^2)\cos\theta + 2R_1R_2\sin\theta\,,$$

which we call "quantum convex" combination of $\Re\mathcal{B} = R_1^2 - R_2^2$ and $\Im\mathcal{B} = 2R_1R_2$. The word "quantum" honors the fact that coefficients of the combination $\alpha = \cos\theta, \beta = \sin\theta$ satisfy the condition $\alpha^2 + \beta^2 = 1$.

Theorem 6. *1. For any matrix A we have the estimate from above*

$$\|T_A\|_p \le \|A\|(p^*-1)\,.$$

2. For any $\theta\in[0,2\pi)$ we have $\|T_{A^\theta}\|_p = p^*-1$.

3. In particular ($\theta=0$) $\|R_1^2-R_2^2\| = p^*-1$ and ($\theta=\pi$) $\|2R_1R_2\| = p^*-1$.

Proof. **The estimate from above.**

We introduce real martingales a_2, a_3 as follows

$$a_2(t) - ia_3(t) := \int_0^t \nabla U_{\bar g}(W_s, T-s)\cdot dW_s = U_{\bar g}(W_t, T-t) - U_{\bar g}(0,T)\,. \tag{3.8}$$

Also we introduce real martinagales a_5, a_6 as follows

$$a_5(t) + ia_6(t) := \int_0^t A\nabla U_f(W_s, T-s)\cdot dW_s\,. \tag{3.9}$$

We will need also

$$a_1(t) := U_{|f|^p}(W_t, T-t)\,, \tag{3.10}$$

$$a_4(t) := U_{|g|^q}(W_t, T-t)\,, \tag{3.11}$$

These are martingales too.

To estimate (see (2.13))

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar g(x)\,dx = \lim_{T\to\infty}\mathcal{E}\int_0^T (d\langle a_2, a_5\rangle(t) + d\langle a_3, a_6\rangle(t))\,dt \tag{3.12}$$

we will use "a monotone functional" approach.

We have B from Theorem 3. Consider martingale $a=(a_1,a_2,a_3,a_4,a_5,a_6)$, and put

$$\Phi(t) = \mathcal{E}\{B(a(t))\}\,.$$

Now B satisfies property I). Summing up property I) for $\alpha := \alpha^k, k=1,..,d$ and applying (3.6) we get

$$-\Phi'(t) \ge \mathcal{E}(d\langle \alpha_2, \alpha_5\rangle(t) + d\langle\alpha_3,\alpha_6\rangle(t)) \tag{3.13}$$

Let us combine (3.12) and (3.13) to obtain

$$\int_{\mathbb{R}^d} T_A\, f(x)\bar g(x)\,dx = \lim_{T\to\infty}\int_0^T(-\Phi'(t))\,dt \le \Phi(0)\,.$$

Now we use property II) to see that

$$\Phi(0) \le \lim_{T\to\infty}(p^*-1)\mathcal{E}a_1(0)^{1/p}a_4(0)^{1/q} = (p^*-1)\lim_{T\to\infty}(2\pi T)^{\frac{d}{2}}\mathbb{E}a_1(0)^{1/p}a_4(0)^{1/q} \le$$

$$\|A\|(p^*-1)\lim_{T\to\infty}(2\pi T)^{\frac{d}{2}}(U_{|f|^p}(0,T))^{1/p}(U_{|g|^q}(0,T))^{1/q}=(p^*-1)\|f\|_p\|g\|_q\,.$$

Estimate from above is completely finished. The estimate 2) for quantum combinations $T_{A^\theta}=(R_1^2-R_2^2)\cos\theta+2R_1R_2\sin\theta$ follows from the fact that $\|A^\theta\|=1$.

The estimate from below on $\|(R_1^2-R_2^2)\cos\theta+2R_1R_2\sin\theta\|=\|T_{A^\theta}\|$.

In Section 11 we will prove that

$$\|T_{A^\theta}\|\ge p^*-1\,,\ \|T_A\|\ge p^*-1\,.$$

□

A famous Iwaniec-Gehring conjecture claims that for the Ahlfors-Beurling operator the same estimate p^*-1 holds as for its real and imaginary parts in Theorem 5.

Conjecture 7.

$$\|\mathcal{B}\|=\|R_1^2-R_2^2+2iR_1R_2\|=p^*-1\,.$$

By theorem 6 (see also Section 11) we know only

$$p^*-1\le\|\mathcal{B}\|\le 2(p^*-1)\,. \tag{3.14}$$

We will give better estimates for $\|\mathcal{B}\|$ in Section 6.

4. Heat flow and Bellman function.

Here we illustrate the heat flow approach to estimation of operators $T_A=\sum_{k,l=1}^d a_{kl}R_kR_l$. The first claim of Theorem 6: $\|T_A\|\le||A||(p^*-1)$ will be proved by the heat flow method without the use of stochstic integration, without Itô's formula, without any martingales. This will be a *a purely analytic* proof. However, the attentive reader will notice that actually what we are doing in this section is getting read of martingales $A\star F$. We will be using (but implicitly, under disguise) the heat martingales. Heat martingales are stochastic processes $U_f(Z_t)$, where $U_f(\cdot,\cdot)$ is the heat extension of function f on $\mathbb{R}^d$ into the half-space $\mathbb{R}^{d+1}_+$, and $Z_t=(W_t,T_t)$ is the so-called "space-time" Brownian motion.

Now formula (2.6) is our starting point:

$$(T_A\,f,\,g)=\int_{\mathbb{R}^{d+1}_+}A\nabla_xU_f(x,t)\cdot\nabla_xU_g(x,t)\,dxdt\,, \tag{4.1}$$

where $\cdot$ denotes the scalar product in $\mathbb{C}^d$.

We again use the same function B from Theorem 3. It satisfies properties I) and II).

Let f, g be two test functions (smooth with compact support). The heat flow method consists of introducing

$$\Phi(t) := \int_{\mathbb{R}^d} B(U_{|f|^p}(x,t), U_{\Re f}(x,t), U_{\Re g}(x,t), U_{\Im g}(x,t), U_{|g|^q}(x,t), U_{\Re f}(x,t), U_{\Im f}(x,t))\, dx\,.$$

We will see that $\Phi(t)$ is decreasing when $t \to +\infty$, so the heat flow dissipates the energy. Th enargy balance will give us this claim of Theorem 6: $\|T_A\| \le \|A\|(p^*-1)$.

To do that let us first use the notation

$$a(x,t) = (U_{|f|^p}(x,t), U_{\Re f}(x,t), U_{\Re g}(x,t), U_{\Im g}(x,t), U_{|g|^q}(x,t), U_{\Re f}(x,t), U_{\Im f}(x,t))$$

and write the chain of inequalities:

$$\Phi(0) \ge \int_0^{+\infty} (-\Phi'(t))\, dt = \int_0^{+\infty} \int_{\mathbb{R}^d} \frac{\partial}{\partial t} B(a(x,t))\, dxdt =$$
$$\int_0^{+\infty} \int_{\mathbb{R}^d} (\frac{\partial}{\partial t} - \frac{1}{2}\Delta) B(a(x,t))\, dxdt\,. \tag{4.2}$$

The last equality being the manifestation of the fact that

$$\int_{\mathbb{R}^d} \Delta B(a(x,t))\, dx = 0\,.$$

In fact, we need to see (after integration by parts) that $\nabla B(a(x,t)) \cdot \frac{\partial}{\partial n} a(x,t)$ tend to zero when $|x| = R\,, R \to \infty$. This is so because $a(x,t)$, $\frac{\partial}{\partial n} a(x,t)$ go to zero very fast when x goes to infinity (after all $f, g, |f|^p, |g|^q$ are functions with compact support), and $B(0) = 0$, moreover, $|\nabla B(a)| \le C\|a\|^{-\delta}$ with $\delta < 1$. Notice also that by property II) of B we have

$$\Phi(0) \le$$
$$(p^*-1)\int_{\mathbb{R}^d} (U_{|f|^p}(x,0))^{1/p} (U_{|g|^q}(x,0))^{1/q}\, dx \le (p^*-1)\|f\|_p \|g\|_q\,. \tag{4.3}$$

Let $\frac{d^2 B}{d\,a^2}$ denote the Hessian matrix of B.

Lemma 8.

$$(\frac{\partial}{\partial t} - \frac{1}{2}\Delta) B(a(x,t)) = -\frac{1}{2}\sum_{k=1}^{d} \langle \frac{d^2 B}{d\,a^2} \frac{\partial a}{\partial x_k}, \frac{\partial a}{\partial x_k} \rangle\,,$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{C}^6$.

Taking this lemma for granted for a while we continue by using property I) of B to have

$$-\frac{1}{2}\sum_{k=1}^{d}\langle\frac{d^2B}{d\,a^2}\frac{\partial a}{\partial x_k},\frac{\partial a}{\partial x_k}\rangle \geq \sum_{k=1}^{d}(|\frac{\partial\Re U_g}{\partial x_k}||\frac{\partial\Re U_f}{\partial x_k}|+|\frac{\partial\Im U_g}{\partial x_k}||\frac{\partial\Im U_g}{\partial x_k}|)\,. \tag{4.4}$$

Now the reader should return to (4.1). If we were able to estimate from above $A\nabla_x U_f(x,t)\cdot\nabla_x U_g(x,t)$ by $\|A\|\sum_{k=1}^{d}(|\frac{\partial\Re U_g}{\partial x_k}||\frac{\partial\Re U_f}{\partial x_k}|+|\frac{\partial\Im U_g}{\partial x_k}||\frac{\partial\Im U_g}{\partial x_k}|)$, we would obtain

$$|(T_A\,f,g)|\leq\Phi(0)\,, \tag{4.5}$$

and in conjunction with (4.3) it would give the assertion 1) of Theorem 6.

Of course, the estimate from above

$$A\nabla_x U_f(x,t)\cdot\nabla_x U_g(x,t)\;\text{ by }\;\|A\|\sum_{k=1}^{d}(|\frac{\partial\Re U_g}{\partial x_k}||\frac{\partial\Re U_f}{\partial x_k}|+|\frac{\partial\Im U_g}{\partial x_k}||\frac{\partial\Im U_g}{\partial x_k}|)$$

is false in general even for real-valued f,g. Just $|\sum_{k,l=1}^{d}a_{kl}x_k y_l|$ cannot be bounded by $\|A\|\sum_{k=1}^{d}|x_k||y_k|$.

As we cannot improve the estimate from above, may be we should strengthen the estimate (4.4) from below. In fact, suppose we can prove an estimate of type (4.4) but much stronger:

$$-\frac{1}{2}\sum_{k=1}^{d}\langle\frac{d^2B}{d\,a^2}\frac{\partial a}{\partial x_k},\frac{\partial a}{\partial x_k}\rangle \geq$$

$$(\sum_{k=1}^{d}|\frac{\partial\Re U_g}{\partial x_k}|^2+|\frac{\partial\Im U_g}{\partial x_k}|^2)^{\frac{1}{2}}(\sum_{k=1}^{d}|\frac{\partial\Re U_f}{\partial x_k}|^2+|\frac{\partial\Im U_f}{\partial x_k}|^2)^{\frac{1}{2}} =: \|\nabla U_g\|\|\nabla U_f\|\,. \tag{4.6}$$

Suppose (4.6) is already proved. Then we combine (4.1) with $|A\nabla_x U_f(x,t)\cdot\nabla_x U_g(x,t)|\leq\|A\|\|\nabla U_f\|\|\nabla U_g\|$ to have

$$|(T_A\,f,\,g)|\leq\|A\|\int_0^{+\infty}\int_{\mathbb{R}^d}\|\nabla_x U_f(x,t)\|\,\|\nabla_x U_g(x,t)\|\,dxdt\,, \tag{4.7}$$

Combine this with (4.6) and Lemma 8 to get

$$|(T_A\,f,\,g)|\leq\|A\|\int_0^{+\infty}\int_{\mathbb{R}^d}-\frac{1}{2}\sum_{k=1}^{d}\langle\frac{d^2B}{d\,a^2}\frac{\partial a}{\partial x_k},\frac{\partial a}{\partial x_k}\rangle\,dxdt=$$

$$\|A\|\int_0^{+\infty}\int_{\mathbb{R}^d}(\frac{\partial}{\partial t}-\frac{1}{2}\Delta)B(a(x,t))\,dxdt\leq\|A\|\Phi(0)\,. \tag{4.8}$$

The last assertion follows from our chain of inequalities preceeding (4.3). Now we use (4.3) and we get from it and (4.8) that

$$|(T_A\, f, g)| \le \|A\| \|f\|_p \|g\|_q\,.$$

4.1. **Bilinear embedding theorem.** We are ready to formulate the following **bilinear embedding** theorem:

Theorem 9.

$$\int_{\mathbb{R}^{d+1}_+} \Big(\sum_{k=1}^d \Big(\Big|\frac{\partial \Re U_g}{\partial x_k}\Big|^2 + \Big|\frac{\partial \Im U_g}{\partial x_k}\Big|^2\Big)\Big)^{1/2} \Big(\sum_{k=1}^d \Big(\Big|\frac{\partial \Re U_f}{\partial x_k}\Big|^2 + \Big|\frac{\partial \Im U_f}{\partial x_k}\Big|^2\Big)\Big)^{1/2} dxdt \le (p^* - 1)\|f\|_p \|g\|_q\,.$$

Proof. Integrating (4.6) over $\mathbb{R}^{d+1}_+$ and using Lemm 8 we can apply (4.2) and then (4.3). This proves the theorem.

□

4.2. **The proof of** (4.6)**.** To prove (4.6) one needs to show that

$$-\sum_{k=1}^d \langle \frac{d^2 B}{d\,a^2}\alpha^k, \alpha^k\rangle \ge 2\,(\sum_{k=1}^d \{(\alpha_2^k)^2 + (\alpha_3^k)^2\})^{\frac12} (\sum_{k=1}^d \{(\alpha_5^k)^2 + (\alpha_6^k)^2\})^{\frac12}\,. \tag{4.9}$$

We have only propert I), which gives that for every individual $k = 1, ..., d$ we have

$$-\langle \frac{d^2 B}{d\,a^2}\alpha^k, \alpha^k\rangle \ge 2\,((\alpha_2^k)^2 + (\alpha_3^k)^2)^{\frac12} ((\alpha_5^k)^2 + (\alpha_6^k)^2)^{\frac12}\,. \tag{4.10}$$

Amazingly the latter implies the former (going “against” Cauchy inequality).

Self-improvment lemma for three quadratic forms.

Here is a lemma of Marcus about matrix pencils [M]. Matrix pencil is $D\lambda^2 + A\lambda + C$, where A, C, D are matrices.

Lemma 10. *Let $p(\lambda) = D\lambda^2 - A\lambda + C$ be a matrix pencil, where A, D, C are non-negatively defined $d \times d$ matrices. Suppose that for every vector $h \in \mathbb{C}^d$ there exists $\lambda = \lambda(h)$ such that $\langle p(\lambda)h, h\rangle \le 0$. Then there exists a universal λ_0 such that for all $h \in \mathbb{C}^d$ we have $\langle p(\lambda_0)h, h\rangle \le 0$.*

We will prove this lemma below in a much higher generality. Now let us use it.

Corollary 11. *Let A, D, C be three non-negatively defined matrices such that*

$$\langle Ah, h\rangle \geq 2\,\langle Dh, h\rangle^{\frac{1}{2}}\,\langle Ch, h\rangle^{\frac{1}{2}}\,. \tag{4.11}$$

Then

$$trace\,A \geq 2\,(trace\,D)^{\frac{1}{2}}\,(trace\,C)^{\frac{1}{2}}\,. \tag{4.12}$$

Proof. Let $\{\alpha^k\}_{k=1}^d$ denote the orthonormal basis of $\mathbb{C}^d$. Lemma claims that independently of vector α there exists $\tau > 0$ such that

$$\langle A\alpha, \alpha\rangle \geq \tau\langle D\alpha, \alpha\rangle + \frac{1}{\tau}\langle C\alpha, \alpha\rangle\,.$$

Applying this to all vectors $\alpha = \alpha^k$, $k = 1, .., d$ and adding in k we obtain

$$\text{trace}\,A \geq \tau\text{trace}\,D + \frac{1}{\tau}\text{trace}\,C\,.$$

This implies (4.12) by arithmetic mean/geometric mean inequality. □

Now we want to use Corollary 11 to prove (4.9). To do that consider a real matrix $\mathcal{A}$ with columns $\alpha^1 = (\alpha_1^1, \ldots, \alpha_6^1)^T$,...,$\alpha^d = (\alpha_1^d, \ldots, \alpha_6^d)^T$. Put $A = -\mathcal{A}^* \frac{d^2 B}{d\,a^2}\mathcal{A}$, then we see that

$$\text{trace}\,A = -\sum_{k=1}^{d}\langle\frac{d^2 B}{d\,a^2}\alpha^k, \alpha^k\rangle\,.$$

Let d be a 6×6 matrix for which all elements are 0 except $d_{22} = 1, d_{33} = 1$, and c such that all elements are 0 except $c_{55} = 1, c_{66} = 1$. Consider $D = \mathcal{A}^* d\mathcal{A}$, $C = \mathcal{A}^* c\mathcal{A}$.

Then

$$\text{trace}\,D = \sum_{k=1}^{d}\{(\alpha_2^k)^2 + (\alpha_3^k)^2\}\,,$$

$$\text{trace}\,C = \sum_{k=1}^{d}\{(\alpha_5^k)^2 + (\alpha_6^k)^2\}\,.$$

We have inequality (4.11): $\langle A\alpha, \alpha\rangle \geq 2\langle D\alpha, \alpha\rangle^{\frac{1}{2}}\langle C\alpha, \alpha\rangle^{\frac{1}{2}}$ for all vectors $\alpha \in \mathbb{C}^d$. In fact, it is just (4.10). Then Corollary claims (4.9).

The proof of Theorem 6 will be finished when we prove Lemma 8 and Lemma 10.

4.3. **Proofs of Lemma 8 and of the three quadratic forms Lemma 10.** Lemma 8 is very easy. It is just the chain rule.

Proof. The chain rule obviously gives

$$(\frac{\partial}{\partial t}-\frac{1}{2}\Delta)B(a(x,t))=\langle\nabla B(a),\frac{\partial a}{\partial t}\rangle-\langle(\nabla B)\circ a,\frac{1}{2}\Delta_x a\rangle+$$

$$-\frac{1}{2}\sum_{k=1}^{d}\langle\frac{d^2B}{d\,a^2}\frac{\partial a}{\partial x_k},\frac{\partial a}{\partial x_k}\rangle\,,$$

First two terms disappear just because $\frac{\partial a_i}{\partial t}-\frac{1}{2}\Delta a_i=0,\ i=1,...,6.$ □

Lemma 10 can be proved in a much more general situation than we need here [DTV]. It is just a particular case of the following Theorem.

Theorem 12. *Let $\sigma_0,\sigma_1,\sigma_2$ be non-negative non-zero quadratic forms on a vector space $\mathcal{V}$ satisfying*

$$\sigma_0[x]\geqslant 2\sqrt{\sigma_1[x]\cdot\sigma_2[x]}\,,\qquad\forall x\in\mathcal{V}. \tag{4.13}$$

Then there exists a constant $\alpha>0$ such that

$$\sigma_0[x]\geqslant\alpha\sigma_1[x]+\frac{1}{\alpha}\sigma_2[x]\,,\qquad\forall x\in\mathcal{V}. \tag{4.14}$$

The condition (4.14) clearly implies (4.13). Also, we emphasize that the statement is valid for arbitrary vector spaces, i.e. regardless of them admitting properties such as completeness, separability, finite or infinite dimension, real or complex field of scalars, inner product etc.

Proof. Consider the family of quadratic forms σ^s, $s\in(0,\infty)$

$$\sigma^s:=\sigma_0-s\sigma_1-s^{-1}\sigma_2.$$

Assume that the theorem is not true, i.e. that for all s the form σ^s is not non-negative to get a contradiction.

First of all let us notice, that if for some $x\in\mathcal{V}$ we have $\sigma_2[x]/\sigma_1[x]=s^2$ then

$$s\sigma_1[x]+s^{-1}\sigma_2[x]=2\sqrt{\sigma_1[x]\cdot\sigma_2[x]}.$$

Therefore, if we find $s>0$ such that

$$\sigma^s[x]<0$$

for some x satisfying $\sigma_2[x]/\sigma_1[x]=s^2$, we get a contradiction with (4.13).

Consider the set $S \subset (0, \infty) \times (0, \infty)$ consisting of all pairs (s, t) such that $\sigma^s[x] < 0$ for some x satisfying $\sigma_2[x]/\sigma_1[x] = t^2$. We want to show that $(\alpha, \alpha) \in S$ for some $\alpha > 0$, which gives us the contradiction.

Since all forms are non-zero, there exist vectors x_k, $k = 1, 2$ such that $\sigma_k[x_k] > 0$, $k = 1, 2$. Therefore there exists a linear combination $x = \alpha x_1 + \beta x_2$ such that $\sigma_1[x], \sigma_2[x] > 0$ (to see that one only needs to consider forms on two-dimensional space $\mathcal{L}\{x_1, x_2\}$). One concludes that $\sigma^s[x] < 0$ for all sufficiently large and for all sufficiently small s. So, if $t_0 := \sqrt{\sigma_2[x]/\sigma_1[x]}$, the points (s, t_0) belong to S for all sufficiently small and for all sufficiently large s. Thus S has points on both sides of the line $s = t_0$. So, if we prove that the set S is connected, it must contain a point (α, α), which gives us the desired contradiction.

We now prove the following properties of the set S:

(i) For any $(s_0, t_0) \in S$ we have $(s, t_0) \in S$ for all s in a small neighborhood of s_0.

(ii) Projection of S onto the s-axis is the whole ray $(0, \infty)$.

(iii) For any $s \in (0, \infty)$ the set $\{t : (s, t) \in S\}$ is an interval.

(iv) S is connected.

Property (i) follows immediately from the continuity of the function $s \mapsto \sigma^s[x]$ (x is fixed).

Property (ii) is just our assumption that σ^s is never positive semi-definite.

Property (iii) requires some work. Let s be fixed. Suppose that $(s, t_k) \in S$, $k = 1, 2$, i.e. that there exist vectors $x_1, x_2 \in \mathcal{V}$ such that $\sigma^s[x_k] < 0$ and $t_k = \sqrt{\sigma_2[x_k]/\sigma_1[x_k]}$, $k = 1, 2$. Consider the (real) subspace $\mathcal{E} \subset \mathcal{V}$, $\mathcal{E} = \operatorname{span}_{\text{real}}\{x_1, x_2\}$, and let us restrict all quadratic forms onto $\mathcal{E}$.

For a vector $x \in \mathcal{E}$ satisfying $\sigma^s[x] < 0$ define

$$\tau(x) := \sqrt{\sigma_2[x]/\sigma_1[x]}.$$

Notice, that by the definition of σ^s for any x satisfying $\sigma^s[x] < 0$ both $\sigma_1[x]$ and $\sigma_2[x]$ cannot be simultaneously 0, so $\tau : \{x \in \mathcal{E} : \sigma^s[x] < 0\} \to [0, \infty]$ is a well defined continuous map (we are allowing $\tau(x) = +\infty$).

Since $\sigma^s[x_k] < 0$ the quadratic form $\sigma^s \bigm| \mathcal{E}$ has either one or two negative squares. In the latter case the set $K = \{x \in \mathcal{E} : \sigma^s[x] < 0\}$ is the whole plane without the origin, so it is connected. In the former case it consists of two connected parts $K = K_1 \cup K_2$, $K_1 = -K_2$. In both cases the set $\tau(K)$ is connected. Indeed, if K

is connected, $\tau(K)$ is a continuous image of a connected set. In the second case, $\tau(K_1)$ is connected and since $\tau(x) = \tau(-x)$ we have $\tau(K_1) = \tau(K_2) = \tau(K)$. So the set $\tau(K)$ contains the whole interval between the points t_1 and t_2.

Since $\tau(K) \cap (0, \infty) \subset \{t : (s,t) \in S\}$ we can conclude that for arbitrary $t_1, t_2 \in \{t : (s,t) \in S\}$, $t_1 < t_2$, the whole interval $[t_1, t_2]$ belongs to the set. But that exactly means that the set $\{t : (s,t) \in S\}$ is an interval.

And now let us prove property (iv) (and so the theorem). Suppose we split S into 2 nonempty disjoint relatively open subsets $S = S_1 \cup S_2$. Let P denote the coordinate projection onto the s-axis. Property (i) implies that the sets PS_1, PS_2 are open. Property (ii) implies that $PS_1 \cup PS_2 = (0, \infty)$ so it follows from the connectedness of $(0, \infty)$ that $PS_1 \cap PS_2 \neq \varnothing$.

Therefore for some s there exist t_1, t_2 such that $(s, t_k) \in S_k$, $k = 1, 2$. By property (iii) the whole interval $J = \{(s, \theta t_1 + (1-\theta)t_2) : \theta \in [0,1]\}$ belongs to S. Therefore J can be represented as a union $J = (J \cap S_1) \cup (J \cap S_2)$ of disjoint nonempty relatively open subsets, which is impossible.

□

5. Better estimates of $\|\mathcal{B}\|$ via heat flow.

We will use Theorem 6 to improve the estimate for Ahlfors-Beurling operator. So far we know only $\|\mathcal{B}\| \leq 2(p^* - 1)$. Notice that it is sufficinet to give a better estimate only for $p \geq 2$ because then we use a simple fact that $\mathcal{B}^*$ looks exactly like $\mathcal{B}$ to get the estimate for $p < 2$.

The next theorem is not the sharpest we know. The ultimate result of Banuelos-Janakiraman [BaJa1] will be explained in the next Section 6. But the next result however illustrates the heat flow method and it is asymptotically as sharp as [BaJa1], it gives $\|\mathcal{B}\| \leq (\sqrt{2} + 0(1))(p^* - 1)$, when $p \to \infty$.

Moreover, in the course of proving Theorem 16 below we will prove an estimate on L^p norm of

$$[|(R_1^2 - R_2^2)u|^2 + (R_1^2 - R_2^2)v|^2 + |2R_1R_2u|^2 + |2R_1R_2v|^2]^{\frac{1}{2}},$$

via the L^p norm of $[u^2 + v^2]^{\frac{1}{2}}$ which we consider interesting.

5.1. **A theorem of Marcienkiewicz and Zygmund and its extension.** Let $\mathcal{R}^p$ denotes the class of operators mapping **real valued** $L^p(\mu)$ to **real valued** $L^p(\nu)$ and bounded. Every such operator can be obviously *complexified* and called

T_c. Moreover, given a real separable Hilbert space H it can be *tenzorized* to $T_H := T \otimes Id_H$ that acts from H **valued** $L^p(H,\mu)$ to (hopefully) H **valued** $L^p(H,\nu)$. Denoting l^2_n the n-dimensional space of square summable real sequences (of course l^2 stands for l^2_n with $n=\infty$) we see that T_c and $T_{l^2_2}$ are isometrically isomorphic. A simple but important theorem of Marcienkiewicz and Zygmund claims that tenzorization does not increase the norm.

Theorem 13.

$$\|T_H\|_p = \|T\|_p\,.$$

In particular, $\|T_c\|_p = \|T\|_p$.

Proof. We need only to prove inequality $\|T_H\|_p \le \|T\|_p$. because the converse is obvious. Let H be $n \le \infty$ dimensional separable real Hilbert space. Let $\{\xi_i\}_{i=1}^n$ be sequence of independent gaussian real random varibles such that $E\xi_i = 0$, $E(\xi_i)^2 = 1$. Fix an element $a = \{a_k\}_{k=1}^n \in l^2_n$. Then we have

$$E(|\sum_{i=1}^n a_i\xi_i|^p) = A_p(\sum_{i=1}^n a_i^2)^{\frac{p}{2}}\,, \tag{5.1}$$

where $A_p := \mathbb{E}|\xi_i|^p$. This obvious because the distribution of $\{\xi_i\}_{i=1}^n$ is rotation invariant and so we can think that $a = (1,0,0,...)$.

Now fix a basis in H. Then an element f of $L^p(H,d\mu)$ can be viewed as $f = \{f_k(x)\}_{k=1}^n$. Consider a new (random) element of scalar real $L^p(d\mu)$: $\sum_k \xi_k f_k(x)$. Then $\mathbb{E}\|\sum_k \xi_k f_k(x)\|^p_{L^p(\mu)} = A_p\int(\sum_{i=1}^n f_i^2(x))^{\frac{p}{2}}\,d\mu(x) = A_p\|f\|^p_{L^p(H,d\mu)}$ by (5.1). Similarly

$$\mathbb{E}\|\sum_k \xi_k Tf_k(x)\|^p_{L^p(\mu)} = A_p\int(\sum_{i=1}^n (Tf_i)^2(x))^{\frac{p}{2}}\,d\mu(x) = A_p\|Tf\|^p_{L^p(H,d\mu)}\,.$$

But

$$\|\sum_k \xi_k Tf_k(x)\|^p_{L^p(\mu)} = \|T(\sum_k \xi_k f_k(x))\|^p_{L^p(\mu)} \le \|T\|_p^p\,\|\sum_k \xi_k f_k(x)\|^p_{L^p(\mu)}\,.$$

Comparing this we see that

$$A_p\|Tf\|^p_{L^p(H,d\mu)} \le A_p\|T\|_p^p\|f\|^p_{L^p(H,d\mu)}\,.$$

Cancel A_p to get the desired inequality. □

In what follows we will need the extension of this theorem to operators from $L^p(\mu)$ to $L^p(H,\nu)$. Here H as before is a real separable Hilbert space (actually we will use only $H = l^2_2$), and $L^p(\mu)$ is again real valued. Class $\mathcal{R}^p_H$ denotes all the bounded linear operators from $L^p(\mu)$ to $L^p(H,\nu)$. It can be obviously extended to T_c, or more generally to operator $T_{l^2_n} = T \otimes Id_{l^2_n}$ from $L^p(l^2_n,\mu)$ to $L^p(l^2_n(H),\nu)$ (with the ususal convention $n \le \infty$, and $l^2_n = l^2$ for $n=\infty$). We just put $T_{l^2_n}(\{f_k(x)\}_{k=1}^n) = \{Tf_k(x)\}_{k=1}^n$. Obviously T_c is isometrically isomorphic to $T_{l^2_2}$ again. We ask the same question $\|T_{l^2_n}\|_p = \|T\|_p$? Or simply $\|T_c\|_p = \|T\|_p$? We are grateful to A. Alexandrov who showed the following result.

Theorem 14. *1. If $0<p\le 2$ then $\|T_{l^2_n}\|_p = \|T\|_p$, and, in particular, $\|T_c\|_p = \|T\|_p$.*

2. If $p>2$ then there are bounded operators T such that $\|T_c\|_p > \|T\|_p$.

Proof. 1. Again we need only $\|T_{l^2_n}\|_p \le \|T\|_p$. Let ξ_i, $i=1,...,n$, and $A_p = E|\xi_i|^p$ be as before. Let $a=\{a_i\}_{i=1}^n \in l^2_n(H)$. Then for $0<p\le 2$

$$E(\|\sum_{i=1}^n a_i\xi_i\|_H^p) \ge A_p(\sum_{i=1}^n \|a_i\|_H^2)^{\frac{p}{2}}\,, \tag{5.2}$$

To prove let us denote by $\{e_j\}$ an orthonormal basis of H. Then the LHS is

$$E(\sum_{j=1}^\infty(\sum_{i=1}^n (a_i,e_j)\xi_i)^2)^{\frac{p}{2}} = E(\sum \phi_j)^{\frac{p}{2}}\,.$$

We denoted $\phi_j := (\sum_{i=1}^n (a_i,e_j)\xi_i)^2 \ge 0$. We want to use the inverse Cauchy inequality. Let $0<\alpha\le 1$, $L^\alpha(\Omega,dP)$ is a space with probability measure, and $\phi_i \ge 0$ be measurable functions. Then

$$\|\sum_j \phi_j\|_\alpha \ge \sum_j \|\phi_j\|_\alpha\,. \tag{5.3}$$

If we use this inequality for $\alpha = \frac{p}{2}$ we get

$$E(\sum_{j=1}^\infty(\sum_{i=1}^n (a_i,e_j)\xi_i)^2)^{\frac{p}{2}} \ge (\sum_j (E((\phi_j)^{\frac{p}{2}}))^{\frac{2}{p}})^{\frac{p}{2}}\,.$$

Using the fact that $\phi_j^{\frac{p}{2}} = (\sum_{i=1}^n (a_i,e_j)\xi_i)^p$ and (5.1) we get that

$$(\sum_j E((\phi_j)^{\frac{p}{2}}))^{\frac{2}{p}})^{\frac{p}{2}} = A_p(\sum_j\sum_i (a_i,e_j)^2)^{\frac{p}{2}} = A_p(\sum_i \|a_i\|_H^2)^{\frac{p}{2}}\,.$$

So (5.2) is proved. To finish the first part of the theorem consider $f\in L^p(l^2_n,\mu)$, $f=\{f_i\}_{i=1}^n$ and write $a_i(x) = (Tf_i)(x) \in H$. We know that

$$\|\sum_{i=1}^n a_i(x)\xi_i\|^p_{L^p(H,\nu)} \le \|T\|_p^p \|\sum_{i=1}^n f_i\xi_i\|_{L^p(\mu)}$$

Apply E and (5.2). Then (5.2) gives

$$A_p\|Tf\|^p_{L^p(l^2_n(H),\nu)} = A_p\|(\sum_i \|a_i\|^2_H)^{\frac{1}{2}}\|^p_{L^p(\nu)} \le E\|\sum_{i=1}^n a_i(x)\xi_i\|^p_{L^p(H,\nu)} \le$$

$$\|T\|^p_p E\|\sum_{i=1}^n f_i\xi_i\|^p_{L^p(\mu)} = A_p\|T\|^p_p\|(\sum_{i=1}^n f_i^2)^{\frac{1}{2}}\|^p_{L^p(\mu)} = A_p\|T\|^p_p\|f\|^p_{L^p(l^2_n,\mu)}\,.$$

We cancel A_p to get the desired inequality $\|T_{l^2_n}\|_p \le \|T\|_p$. The first part of the theorem is proved.

2. Let $p > 2$. Consider $d\mu = \frac{1}{\pi}e^{-(x^2+y^2)}dxdy$, ν being just a point mass. Operator $T : L^p(\mu) \to L^p(l^2_2, \nu) = l^2_2$ is given by

$$Tf = (\int_{\mathbb{R}^2} x\, f(x,y)\, d\mu(x,y), \int_{\mathbb{R}^2} y\, f(x,y)\, d\mu(x,y)) \in l^2_2 = \mathbb{R}^2\,.$$

Obviously (with $q : 1/q + 1/p = 1$)

$$\sup_{(a,b)\in\mathbb{R}^2: a^2+b^2=1}\ \sup_{\|f\|_{L^p(\mu)}\le 1} \int_{\mathbb{R}^2} (ax+by) f(x,y)\, d\mu(x,y) = (\int_{\mathbb{R}^2} |x|^q\, d\mu(x,y))^{\frac{1}{q}} =: c_q^{\frac{1}{q}}\,.$$

because of rotation invariance of μ. The latter number is the norm of T. Let us check the norm of T_c. Of course

$$\|T_c\|_p \ge \sup_{\|f\|_{L^p_c(\mu)}\le 1} \int_{\mathbb{R}^2} \left(\frac{1}{\sqrt{2}}x + \frac{i}{\sqrt{2}}y\right) f(x,y)\, d\mu(x,y) =$$

$$(\int_{\mathbb{R}^2} \left|\frac{x+iy}{\sqrt{2}}\right|^q d\mu(x,y))^{\frac{1}{q}} =: C_q^{\frac{1}{q}}\,.$$

We want to show that $C_q > c_q$. Notice that

$$2c_q = \int_{\mathbb{R}^2} (|x|^q + |y|^q)\, d\mu(x,y)\,,$$

$$2C_q = \int_{\mathbb{R}^2} \left(\left|\frac{x+iy}{\sqrt{2}}\right|^q + \left|\frac{x-iy}{\sqrt{2}}\right|^q\right) d\mu(x,y)\,.$$

So we need to check an elementary inequality

$$|x|^q + |y|^q < 2^{1-\frac{q}{2}}(x^2+y^2)^{\frac{q}{2}} \text{ for Lebesgue a. e. } (x,y)\,.$$

By homogenuity it is the same as $|\cos\theta|^q + |\sin\theta|^q < 2^{1-\frac{q}{2}}$ for $0 < q < 2$ for a.e. $\theta \in [0, 2\pi)$. But for every fixed $\theta \ne \pm\frac{\pi}{4}, \pm\frac{3\pi}{4}$ we have $|\sqrt{2}\cos\theta|^q + |\sqrt{2}\sin\theta|^q < 2$ as function $\phi(x) := x^{\frac{q}{2}} + (2-x)^{\frac{q}{2}}$ attains its maximum 2 only at $x = 1$ on the interval $[0, 2]$. We are done.

□

5.2. **Estimate of** $\|(|(R_1^2-R_2^2)u|^2+|2R_1R_2u|^2)^{\frac12}\|_p$**.** Denote

$$\tau_p := \left(\frac{1}{2\pi}\int_0^{2\pi}|\cos\theta|^p\,d\theta\right)^{-1/p}. \tag{5.4}$$

Observe that

$$\lim_{p\to\infty}\tau_p = 1\,.$$

Lemma 15. *Let μ be a positive measure on space X and let A, B be operators acting on $L^p_{real}(X,\mu)$. Denote*

$$C(p) := \max_{\theta\in[0,2\pi)}\|A\cos\theta + B\sin\theta\|_p\,.$$

Then

$$\|\begin{pmatrix}A\\B\end{pmatrix}: L^p_{real}\to L^p(\mathbb{R}^2)\| \le \tau_p\,C(p), \qquad \textit{where} \qquad \textit{where } \tau_p \textit{ is defined in } (5.4)\,.$$

Proof. It almost repeats the proof of the frequently used Marcinkiewicz–Zygmund Theorem (5.1). The trick in the proof of Marcinkiewicz–Zygmund theorem can be reduced to acting by T on on $f\cos\theta + g\sin\theta$.

Now instead of acting by T on $f\cos\theta+g\sin\theta$ we fix $u\in L^p_{real}(\mu)$ and apply to it $A\cos\theta + B\sin\theta$. Take $\omega\in[0,2\pi)$ and write temporarily $a = Au(\omega)$, $b = Bu(\omega)$. Since a and b are real (by the assumption of the lemma), there is $\delta=\delta(\omega)\in[0,2\pi)$ such that

$$(a,b) = \sqrt{a^2+b^2}\,(\cos\delta,\,\sin\delta)\in\mathbb{R}^2\,.$$

It follows that

$$a\cos\theta + b\sin\theta = \sqrt{a^2+b^2}\,\cos(\theta-\delta)$$

or, in other words,

$$(A\cos\theta + B\sin\theta)u(\omega) = (|Au(\omega)|^2+|Bu(\omega)|^2)^{1/2}\cos(\theta-\delta(\omega))$$

for all $\theta\in[0,2\pi)$. Consequently,

$$\int_X(|Au(\omega)|^2+|Bu(\omega)|^2)^{p/2}\,|\cos(\theta-\delta(\omega))|^p\,d\mu(\omega)$$

$$=\int_X|(A\cos\theta+B\sin\theta)u(\omega)|^p\,d\mu(\omega)\leqslant C(p)^p\|u\|_p^p\,.$$

Integrate this inequality with respect to the normalized Lebesgue measure $(2\pi)^{-1}d\theta$ on $[0,2\pi)$. We get

$$\left(\int (|Au(\omega)|^2+|Bu(\omega)|^2)^{p/2}d\mu(\omega)\right)^{1/p}\leqslant \tau_p C(p)\|u\|_p\,, \tag{5.5}$$

and Lemma 15 is proved.

□

For $A=R_1^2-R_2^2$ and $B=2R_1R_2$, this lemma gives for real-valued u

$$\left(\int \left(|(R_1^2-R_2^2)u|^2+|2R_1R_2u|^2\right)^{p/2}dxdy\right)^{1/p}\leqslant \frac{(p^*-1)}{\left(\frac{1}{2\pi}\int_0^{2\pi}|\cos\theta|^p\,d\theta\right)^{1/p}}\|u\|_p\,. \tag{5.6}$$

5.3. **Estimate of** $\|(|(R_1^2-R_2^2)u|^2+|2R_1R_2u|^2+|(R_1^2-R_2^2)v|^2+|2R_1R_2v|^2)^{\frac12}\|_p$**.** We want

$$\left(\int \left(|(R_1^2-R_2^2)u|^2+|2R_1R_2u|^2+|(R_1^2-R_2^2)v|^2+|2R_1R_2v|^2\right)^{p/2}dxdy\right)^{1/p}\leqslant$$

$$\left(\frac{1}{2\pi}\int_0^{2\pi}|\cos\theta|^p\,d\theta\right)^{-1/p}(p^*-1)\|(u^2+v^2)^{\frac12}\|_p\,. \tag{5.7}$$

Notice that inequality (5.7) would follow from inequality (5.6) by the complexification of vector operator $T=(R_1^2-R_2^2,2R_1R_2)$ from $L^p(dxdy)$ to $L^p(l_2^2,dxdy)$. (All spaces are real.) It would follow immediately if Theorem 14 would hold for all $p\in(1,\infty)$. Let us explain: again all spaces are real, and let $T:L^p\to L^p(l_2^2)$ be bounded. We would wish to state that then $T:L^p(l_2^2)\to L^p(l_2^2\otimes l_2^2)$ has the same norm. But this is false in general! The second part of Theorem 14 shows that the norm can jump up. However, for $1<p<2$, the norm does not grow and we get (5.3) for $1<p<2$. In particular,

$$\left(\int \left(|(R_1^2-R_2^2)u|^2+|2R_1R_2u|^2+|(R_1^2-R_2^2)v|^2+|2R_1R_2v|^2\right)^{p/2}dxdy\right)^{1/p}\leq$$

$$(\frac{\pi}{2}+o(1))\frac{1}{p-1}\|(u^2+v^2)^{\frac12}\|_p\,,\quad \text{when }\ p\to 1+\ . \tag{5.8}$$

But

$$|\mathcal{B}(u+iv)|=(|(R_1^2-R_2^2)u-2R_1R_2v|^2+|(R_1^2-R_2^2)v+2R_1R_2u|^2)^{\frac12}\,.$$

This and (5.7) immediately imply

$$\left(\int\Big(|\mathcal{B}(u+iv)|\Big)^p\,dxdy\right)^{1/p}\le(\frac{\pi\sqrt{2}}{2}+o(1))\frac{1}{p-1}\|(u^2+v^2)^{\frac12}\|_p\,,\quad\text{when }\ p\to 1+\,. \tag{5.9}$$

We do not know how to get rid of $\frac{\pi}{2}$ in (5.8). But we will show now how to get rid of this constant in (5.9)

Theorem 16. *For $1<p<\infty$ one has*

$$\|\mathcal{B}\|\le\frac{\sqrt{2}}{(\frac{1}{2\pi}\int_0^{2\pi}|\cos\theta|^{p^*}\,d\theta)^{1/p^*}}\,(p^*-1)=(\sqrt{2}+o(1))(p^*-1)\,.$$

Proof. For the length of the proof we adopt the notations: $A:=R_1^2-R_2^2$, $B=2R_1R_2$. Recall that $\mathcal{B}=(R_1+iR_2)^2=A+iB$.

First we write

$$|(A+iB)(u+iv)|^p=(|Au-Bv|^2+|Bu+Av|^2)^{p/2}\,.$$

We will use the same trick as in the proof of Lemma 15, namely, we introduce $\cos\theta,\sin\theta$ as follows. Of course,

$$\Re[(A+iB)(u+iv)e^{-i\theta}]=(Au-Bv)\cos\theta+(Bu+Av)\sin\theta\,,$$

and, therefore, considering a (real-valued) test function ψ, we can write

$$\langle(Au-Bv)\cos\theta+(Bu+Av)\sin\theta,\psi\rangle$$
$$=\Re\int_{\mathbb{R}^2}(A+iB)(u+iv)\psi\,e^{-i\theta}\,dxdy\,.$$

Rewrite the expression by using (2.6). We get

$$\langle(Au-Bv)\cos\theta+(Bu+Av)\sin\theta,\psi\rangle$$
$$=-2\Re e^{-i\theta}\int_{\mathbb{R}^3_+}\left(\frac{\partial}{\partial x_1}+i\frac{\partial}{\partial x_2}\right)(u+iv)\left(\frac{\partial}{\partial x_1}+i\frac{\partial}{\partial x_2}\right)\psi\,.$$

Of course here u,v,ψ are heat extensions of our initial u,v,ψ, that is they satisfy the equation $(\frac{\partial}{\partial t}-\frac12\Delta)(\cdot)=0$ in $\mathbb{R}^3_+$. Let us fix $t>0$ and consider $J(u,v)$, the Jacobian of the map $(x=(x_1,x_2))$ $(u(x,t),v(x,t)):\mathbb{R}^2\to\mathbb{R}^2$. Then our previous equality gives us

$$|\langle(Au-Bv)\cos\theta+(Bu+Av)\sin\theta,\psi\rangle|$$

$$\leqslant \int_{\mathbb{R}^3_+} \left|\left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right)(u+iv)\right| \cdot \left|\left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right)\psi\right|$$

$$\leqslant \int_{\mathbb{R}^3_+} \left(\left|\frac{\partial u}{\partial x_1}\right|^2 + \left|\frac{\partial u}{\partial x_2}\right|^2 + \left|\frac{\partial v}{\partial x_1}\right|^2 + \left|\frac{\partial v}{\partial x_2}\right|^2 - 2\det J(u,v)\right)^{1/2}$$

$$\times \left(\left|\frac{\partial \psi}{\partial x_1}\right|^2 + \left|\frac{\partial \psi}{\partial x_2}\right|^2\right)^{1/2}$$

$$\sqrt{2}\int_{\mathbb{R}^3_+} \left(\left|\frac{\partial u}{\partial x_1}\right|^2 + \left|\frac{\partial u}{\partial x_2}\right|^2 + \left|\frac{\partial v}{\partial x_1}\right|^2 + \left|\frac{\partial v}{\partial x_2}\right|^2\right)^{1/2}$$

$$\times \left(\left|\frac{\partial \psi}{\partial x_1}\right|^2 + \left|\frac{\partial \psi}{\partial x_2}\right|^2\right)^{1/2}$$

$$= \sqrt{2}\int_{\mathbb{R}^3_+} \left(\left|\frac{\partial u + iv}{\partial x_1}\right|^2 + \left|\frac{\partial u + iv}{\partial x_2}\right|^2\right)^{1/2} \left(\left|\frac{\partial \psi}{\partial x_1}\right|^2 + \left|\frac{\partial \psi}{\partial x_2}\right|^2\right)^{1/2},$$

We already established in Theorem 9 that the last integral is bounded by $(p^*-1)\|u+iv\|_p\|\psi\|_q$. Gathering all this together, one obtains

$$|\langle (Au - Bv)\cos\theta + (Bu + Av)\sin\theta, \psi\rangle| \leqslant \sqrt{2}(p^*-1)\|u+iv\|_p\|\psi\|_q\,.$$

And so

$$\|\langle (Au - Bv)\cos\theta + (Bu + Av)\sin\theta\|_p \leqslant \sqrt{2}(p^*-1)\|u+iv\|_p\,. \tag{5.10}$$

Once again we utilize the trick used in the proof of Lemma 15:

$$(Au - Bv)\cos\theta + (Bu + Av)\sin\theta$$
$$= (|Au - Bv|^2 + |Bu + Av|^2)^{1/2}\cos(\theta - \delta(\omega))\,.$$

Therefore, by (5.10),

$$\iint_{\mathbb{R}^2} (|Au - Bv|^2 + |Bu + Av|^2)^{p/2}|\cos(\theta - \delta(\omega))|^p$$
$$= \iint |(Au - Bv)\cos\theta + (Bu + Av)\sin\theta|^p$$
$$\leqslant \sqrt{2}^{\,p}(p^*-1)^p\|u+iv\|_p^p\,.$$

We integrate with respect to normalized Lebesgue measure $dm(\theta)$ as we have already done once before. With τ_p as in (5.4) one completes the proof of Theorem 16.

□

6. Better estimates of $\|\mathcal{B}\|$ via stochastic integrals and Itô's formula.

So far we did not use *local ortogonality* (2), (3.5). To use it we address not to (3.6) but to (3.4):

$$-\Phi'(t) = -\frac{1}{2}\mathcal{E}\sum_{i,j=1}^{6}\frac{\partial^2 B(a(t))}{\partial a_i \partial a_j} d\langle a_i, a_j\rangle(t)\,. \tag{6.1}$$

Notice that *local ortogonality* forces certain terms in the right hand side to disappear. Consider the Hessian matrix $\frac{\partial^2 B(a(t))}{\partial a_i \partial a_j}$ and let us replace entries $\frac{\partial^2 B(a(t))}{\partial a_5 \partial a_6}$, $\frac{\partial^2 B(a(t))}{\partial a_6 \partial a_5}$ by zeroe. The new matrix will be called $N(B)$, its entries will be called $N_{ij}(B)$.

By the force of (3.5) we will have

$$\sum_{i,j=1}^{6}\frac{\partial^2 B(a(t))}{\partial a_i \partial a_j} d\langle a_i, a_j\rangle(t) = \sum_{i,j=1}^{6} N_{ij}(B) d\langle a_i, a_j\rangle(t) = \sum_{k=1}^{d}\langle N(B)a^k, a^k\rangle\,. \tag{6.2}$$

6.1. **Local orthogonality and the Bellster.** Suppose that B satisfies property II) and

I') For an arbitrary $\alpha\in\mathbb{R}^6$ we want to have

$$\langle -N(B)\alpha, \alpha\rangle \geq 2R|\alpha_2||\alpha_5| + 2R|\alpha_3||\alpha_6|\,.$$

Conjecture 17. *Function B satisfying properties I') with $R>1$ and II) exists in the domain $\Omega = \{...\}$.*

This conjecture would have powerful consequences. First let us mix the notations:

$$d\langle X_1, Y_1\rangle = \sum_{k=1}^{d}\alpha_2^k\alpha_5^k\,,\ \ d\langle X_2, Y_2\rangle = \sum_{k=1}^{d}\alpha_3^k\alpha_6^k\,.$$

Using this, summing up property I') for $\alpha := \alpha^k, k = 1,..,d$ and applying (6.2) we get

$$-\Phi'(t) \geq R\mathcal{E}(d\langle X_1, Y_1\rangle + d\langle X_2, Y_2\rangle)$$

and we get from (3.3) a better estimate for the norm of Ahlfors–Beurling operator $\|\mathcal{B}\|$. Recall that $\mathcal{B} = T_A$ in $\mathbb{R}^d, d\geq 2$, and $A = \begin{bmatrix} -1, & i \\ i, & 1\end{bmatrix}$. This A gives rise to *local orthogonality* condition (3.5) and this is why we can use the estimate on

$-N(B)$ which is conjecturally better (see (17)) than the estimate on Hessian matrix $-\frac{d^2B(a}{d^2a}$.

Theorem 18. *If conjecture 17 is satisfied then*

$$\|T_A\| \leq \frac{2}{R}\,(p^*-1)\,.$$

6.2. **Local orthogonality and Burkholder's function.** The material of this section is based on [BaJa1].

Burkholder [Bu1] proposed to consider the following function ($p \geq 2$):

$$\phi(x,y) := p(1-\frac{1}{p})^{p-1}(|y|-(p-1)|x|)(|x|+|y|)^{p-1}\,.$$

He proved the following

Theorem 19. *If* $p \geq 2$*, then*

$$\phi(x,y) \geq |y|^p - (p-1)^p|x|^p\,, \tag{6.3}$$

and one has the following estimate for the Hessian H_ϕ *of* ϕ*:*

$$\langle H_\phi(x,y)\begin{bmatrix} h \\ k \end{bmatrix}, \begin{bmatrix} h \\ k \end{bmatrix}\rangle = p(p-1)(|x|+|y|)^{p-2}(|k|^2-|h|^2)-p(p-1)(p-2)|x|(|x|+|y|)^{p-3}(h+k)^2\,. \tag{6.4}$$

In particular,

$$\langle\frac{d^2\phi(x,y)}{dxdy}\begin{bmatrix} h \\ k \end{bmatrix}, \begin{bmatrix} h \\ k \end{bmatrix}\rangle \leq 0\,,\ \forall\ (h,k)\,, |k| \leq |h|\,. \tag{6.5}$$

If a function $|y|^p - \rho|x|^p$ *has a majorant with property* (6.5)*, then* $\rho \geq (p-1)^p$*.*

This theorem will be proved in Section 10. Its proof there (unlike the original Burkholder's proof) will be based on the solution of Monge-Ampére equation.

Now we want just verify (6.4), and actually we calculate a slightly more difficult thing. For a separable real Hilbert space H let us consider $\Phi(X,Y) = \phi(\|X\|,\|Y\|)$. Let us fix a basis in H and denote the cordinates $X = (x_1,...)$, $Y = (y_1,...)$, $n = \dim H, n \in [1,\infty]$. For $h \in H$ we denote $h' = \mathrm{Proj}_X\, h$, where Proj_X is the ortogonal projection onto the one dimensional space defined by vector X, we also denote $h'' = h - h'$, it is the projection of h onto the complement of X. Symmetrically, given $k \in H$, we denote $k' = \mathrm{Proj}_Y\, k$, where Proj_Y is the ortogonal projection onto the one dimensional space defined by vector Y, we also denote

$k'' = k - k'$, it is the projection of k onto the complement of Y. Then the $2n \times 2n$ Hessian of Φ satisfies

$$\langle H_\Phi(X,Y) \begin{bmatrix} h \\ k \end{bmatrix}, \begin{bmatrix} h \\ k \end{bmatrix} \rangle =$$

$$p(p-1)(\|X\|+\|Y\|)^{p-2}(\|k\|^2-\|h\|^2)-p(p-2)\|Y\|^{-1}(\|X\|+\|Y\|)^{p-1}(\|k\|^2-(k,\frac{Y}{\|Y\|})^2)$$
$$-p(p-1)(p-2)\|X\|(\|X\|+\|Y\|)^{p-3}(\pm\|h'\|+\pm\|k'\|)^2\,,\ \forall h,k\in H\,. \tag{6.6}$$

In fact, by simple calculations

$$\frac{d}{dt}\begin{bmatrix} \|X+th\| \\ \|Y+tk\| \end{bmatrix} = \begin{bmatrix} \frac{\langle X+th,h\rangle+\langle h,X+th\rangle}{2\|X+th\|} \\ \frac{\langle Y+tk,k\rangle+\langle k,Y+tk\rangle}{2\|Y+tk\|} \end{bmatrix}.$$

So

$$v_{h,k} := \frac{d}{dt}\begin{bmatrix} \|X+th\| \\ \|Y+tk\| \end{bmatrix}\Big|_{t=0} = \begin{bmatrix} \pm\|h'\| \\ \pm\|k'\| \end{bmatrix}.$$

Here $\pm\|h'\| = \langle h, \frac{X}{\|X\|}\rangle$ and $\pm\|k'\| = \langle k, \frac{Y}{\|Y\|}\rangle$. Also

$$\frac{d^2}{dt^2}\begin{bmatrix} \|X+th\| \\ \|Y+tk\| \end{bmatrix} = \begin{bmatrix} \frac{\|h\|^2}{\|X+th\|} - \frac{\langle X+th,h\rangle}{\|X+th\|^3} \\ \frac{\|k\|^2}{\|Y+tk\|} - \frac{\langle Y+tk,k\rangle}{\|Y+tk\|^3} \end{bmatrix}.$$

So

$$e_{h,k} := \frac{d^2}{dt^2}\begin{bmatrix} \|X+th\| \\ \|Y+tk\| \end{bmatrix}\Big|_{t=0} = \begin{bmatrix} \frac{\|h''\|^2}{\|X\|} \\ \frac{\|k''\|^2}{\|Y\|} \end{bmatrix}.$$

Recall $e_{h,k}, v_{h,k}$ are vectors in $\mathbb{R}^2$.

Of course $\phi_x(x,y) = -p(p-1)x(x+y)^{p-2}$, $\phi_y(x,y) = p(p-1)y(x+y)^{p-2} - p(p-2)x(x+y)^{p-1}$, and we leave the reader to compute all second derivatives of ϕ.

Below $\langle\cdot,\cdot\rangle$ denote the duality in $\mathbb{R}^2$. Direct calculus of derivatives of $\phi(x,y)$ and the forms of $e_{h,k}, v_{h,k}$ found above show

$$\frac{d^2}{dt^2}\phi(\|X+th\|,\|Y+tk\|)|_{t=0} = \langle\nabla\phi, v_{h,k}\rangle + \langle H_\phi v_{h,k}, v_{h,k}\rangle =$$
$$-p(p-1)(\|X\|+\|Y\|)^{p-2}\|h''\|^2 + p(p-1)(\|X\|+\|Y\|)^{p-2}\|k''\|^2 - p(p-2)(\|X\|+\|Y\|)^{p-1}\frac{\|k''\|^2}{\|Y\|}$$
$$+p(p-1)(\|X\|+\|Y\|)^{p-2}(((\langle k,\frac{Y}{\|Y\|}\rangle)^2 - (\langle h,\frac{X}{\|X\|}\rangle)^2)$$
$$-p(p-1)(p-2)\|X\|(\|X\|+\|Y\|)^{p-3}(\langle h,\frac{X}{\|X\|}\rangle + \langle k,\frac{Y}{\|Y\|}\rangle)^2$$

Now we plug into the third line of this formula $(\langle k, \frac{Y}{\|Y\|}\rangle)^2 = \|k'\|^2 = \|k\|^2 - \|k''\|^2$, $(\langle h, \frac{X}{\|X\|}\rangle)^2 = \|h'\|^2 = \|h\|^2 - \|h''\|^2$ into this formula to get for all $h, k \in H$

$$\langle H_\Phi \begin{bmatrix} h \\ k \end{bmatrix}, \begin{bmatrix} h \\ k \end{bmatrix}\rangle = p(p-1)(\|X\|+\|Y\|)^{p-2}(\|k\|^2-\|h\|^2)-p(p-2)\|Y\|^{-1}(\|X\|+\|Y\|)^{p-1}\|k''\|^2$$

$$-p(p-1)(p-2)\|X\|(\|X\|+\|Y\|)^{p-3}(\langle h, \frac{X}{\|X\|}\rangle + \langle k, \frac{Y}{\|Y\|}\rangle)^2 . \qquad (6.7)$$

which is exactly (6.6).

Let f be a test complex valued functions on the plane. Recall that $Z_t = (W - t, T - t)$ is a space-time Brownian motion on the interval $[0, T]$, W_s is 2-dimensional Brownian motion. Recall also that formula (2.7) can be interpreted as

$$\mathcal{B} f(z) = \lim_{T\to\infty} \mathcal{E}(A \star U_f(Z_T)|\, Z_T := (W_T, 0) = (z, 0)) . \qquad (6.8)$$

Then of course

$$\|\mathcal{B} f\|_p^p \leq \lim_{T\to\infty} \mathcal{E}|A \star U_f(W_t, 0)|^p . \qquad (6.9)$$

Let as above $\Phi(x_1, x_2, y_1, y_2) := \phi(\sqrt{x_1^2 + x_2^2}, \sqrt{y_1^2 + y_2^2})$. Consider now l_2^2-valued martingales $x_{1t} + ix_{2t} := X_t := U_f(Z_t)$, $y_{1t} + iy_{2t} := Y_t = \tau\, A \star U_f(Z_t)$, here $\tau > 0$ will be chosen later. We think about $\mathbb{C}$ as real l_2^2. Introduce

$$a_t := (a_1, a_2, a_3, a_4) := (x_{1t}, x_{2t}, y_{1t}, y_{2t}) .$$

Write Itô's main theorem of (stochastic) calculus ($\langle\cdot,\cdot\rangle$ denote the scalar product in l_2^2, $(\cdot,\cdot)$ denotes the scalar product in l_4^2)

$$|Y_T|^p-(p-1)^p|X_T|^p \leq \Phi(\Re X_T, \Im X_T, \Re Y_T, \Im_T) = \Phi(\Re X_0, \Im X_0, \Re Y_0, \Im_0)+\int_0^T (\nabla\Phi, d\alpha_t)$$

$$+\frac{1}{2}\int_0^T \sum_{i,j=1}^{2} (\Phi_{x_i x_j} d\langle x_{it}, x_{jt}\rangle + \Phi_{x_i y_j} d\langle x_{it}, y_{jt}\rangle + \Phi_{y_i y_j} d\langle y_{it}, y_{jt}) =$$

$$\Phi(\Re X_0, \Im X_0, \Re Y_0, \Im Y_0) + \int_0^T (\nabla\Phi, d\alpha_t) + \frac{1}{2}\int_0^T \sum_{k,l=1}^{4} \frac{d^2\Phi}{da_k\, da_l} d\langle a_k, a_l\rangle . \qquad (6.10)$$

Let us recall what is $d\langle a_k, a_l\rangle$.

First notations.

Now look at (2.11). Notice that for matrix A (and $d = 2$)

$$\overrightarrow{\alpha_{3s}} = \begin{bmatrix} \alpha_3^1 \\ \alpha_3^2 \end{bmatrix} = \tau \begin{bmatrix} (\Re f_x - \Im f_y)(W_s, T - s) \\ (-\Re f_y - \Im f_x)(W_s, T - s) \end{bmatrix} ,$$

$$\overrightarrow{\alpha_{4s}} = \begin{bmatrix} \alpha_4^1 \\ \alpha_4^2 \end{bmatrix} = \tau \begin{bmatrix} (\Re f_y + \Im f_x)(W_s, T-s) \\ (\Re f_x - \Im f_y)(W_s, T-s) \end{bmatrix} . \tag{6.11}$$

Also

$$\overrightarrow{\alpha_{1s}} = \begin{bmatrix} \alpha_1^1 \\ \alpha_1^2 \end{bmatrix} = \tau \begin{bmatrix} (\Re f_x)(W_s, T-s) \\ \Re f_y)(W_s, T-s) \end{bmatrix} ,$$

$$\overrightarrow{\alpha_{2s}} = \begin{bmatrix} \alpha_2^1 \\ \alpha_2^2 \end{bmatrix} = \tau \begin{bmatrix} (\Im f_x)(W_s, T-s) \\ (\Im f_y)(W_s, T-s) \end{bmatrix} . \tag{6.12}$$

Recall that we can write $a_{it}, i = 1, ..., 4$ as stochastic integrals via its stohastic derivatives $\alpha_{is}^k, i = 1, ..., 4,\ k = 1, 2$:

$$a_{it} = \int_0^t (\alpha_{is}^1, \alpha_{is}^2) \cdot dW_s = \int_0^t \overrightarrow{\alpha_{is}} \cdot dW_s \,.$$

Then by Itô's isometry

$$\langle a_{it}, a_{jt} \rangle = \int_0^t \alpha_{is} \cdot \alpha_{js} \, ds \,,$$

where $\cdot$ means the scalar product in l_2^2.

$$\langle a_{3t}, a_{4t} \rangle = \langle y_{1t}, y_{2t} \rangle =$$

$$\int_0^t \tau^2 \Big[(\Re f_x - \Im f_y)(W_s, T-s), \quad (-\Re f_y - \Im f_x)(W_s, T-s) \Big] \cdot \begin{bmatrix} (\Re f_y + \Im f_x)(W_s, T-s) \\ (\Re f_x - \Im f_y)(W_s, T-s) \end{bmatrix} ds = 0$$

In particular, we have (3.5):

$$d\langle a_{3t}, a_{4t} \rangle = d\langle y_{1t}, y_{2t} \rangle = \overrightarrow{\alpha_{3t}} \cdot \overrightarrow{\alpha_{4t}} = 0 \,. \tag{6.13}$$

Simarly, we immediately see that

$$d\langle y_{1t}, y_{1t} \rangle = d\langle a_{3t}, a_{3t} \rangle = |\overrightarrow{\alpha_{3t}}|^2 = d\langle a_{4t}, a_{4t} \rangle = |\overrightarrow{\alpha_{4t}}|^2 = d\langle y_{2t}, y_{2t} \rangle \,. \tag{6.14}$$

Also

$$\begin{aligned} |\overrightarrow{\alpha_{3t}}|^2 + |\overrightarrow{\alpha_{4t}}|^2 &= d\langle y_{1t}, y_{1t} \rangle + d\langle y_{2t}, y_{2t} \rangle \\ = 2\tau^2 (|\nabla U_f|^2 + 2 \det \nabla U_f)(W_t, T-t) &\le 4\tau^2 |\nabla U_f|^2 = \\ 4\tau^2 (d\langle x_{1t}, x_{1t} \rangle + d\langle x_{2t}, x_{2t} \rangle) &= 4\,\tau^2 (|\overrightarrow{\alpha_{1t}}|^2 + |\overrightarrow{\alpha_{2t}}|^2) \,. \end{aligned} \tag{6.15}$$

Here ∇U_f is understood as 2×2 matrix as we have 2 derivatives with respect x and y and f (being complex valued) is understood as a function with values in real l_2^2.

Lemma 20. *Let* $\Phi(a_1,a_2,a_3,a_4) = \phi(\sqrt{a_1^2+a_2^2}, \sqrt{a_3^2+a_4^2})$ *as before. Let* $\alpha^1 = (\alpha_1^1,\alpha_2^1,\alpha_3^1,\alpha_4^1)$ *and* $\alpha^2 = (\alpha_1^2,\alpha_2^2,\alpha_3^2,\alpha_4^2)$ *two vectors such that*

$$\alpha_3^1\alpha_4^1 + \alpha_3^2\alpha_4^2 = 0\,,\ (\alpha_3^1)^2 + (\alpha_3^2)^2 = (\alpha_4^1)^2 + (\alpha_4^2)^2\,. \tag{6.16}$$

Let P, Q *denote projections on the first two and the last two coordinates in* $\mathbb{R}^4$. *Then*

$$\langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^1,\alpha^1\rangle + \langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^2,\alpha^2\rangle \le$$
$$p(p-1)(\|X\|+\|Y\|)^{p-2}(\|Q\alpha^1\|^2+\|Q\alpha^2\|^2-\|P\alpha^1\|^2-\|P\alpha^2\|^2)-$$
$$\frac{1}{2}p(p-2)\|Y\|^{-1}(\|X\|+\|Y\|)^{p-1}(\|Q\alpha^1\|^2+\|Q\alpha^2\|^2)\,.$$

Proof. In fact,

$$\langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^1,\alpha^1\rangle + \langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^2,\alpha^2\rangle \le$$
$$p(p-1)(\|X\|+\|Y\|)^{p-2}(\|Q\alpha^1\|^2+\|Q\alpha^2\|^2-\|P\alpha^1\|^2-\|P\alpha^2\|^2)$$
$$-p(p-2)\|Y\|^{-1}(\|X\|+\|Y\|)^{p-1}(\|Q\alpha^1\|^2+\|Q\alpha^2\|^2-(Q\alpha^1,\frac{Y}{\|Y\|})^2-(Q\alpha^2,\frac{Y}{\|Y\|})^2)$$

Now, using notations $\frac{Y}{\|Y\|} = (y_3,y_4)$ and using (6.16), we get

$$\|Q\alpha^1\|^2+\|Q\alpha^2\|^2-(Q\alpha^1,\frac{Y}{\|Y\|})^2-(Q\alpha^2,\frac{Y}{\|Y\|})^2 = (\alpha_3^1)^2+(\alpha_4^1)^2+(\alpha_3^2)^2+(\alpha_4^2)^2-(\alpha_3^1y_3+\alpha_4^1y_4)^2-$$
$$(\alpha_3^2y_3+\alpha_4^2y_4)^2 = (\alpha_3^1)^2+(\alpha_4^1)^2+(\alpha_3^2)^2+(\alpha_4^2)^2-(\alpha_3^1)^2+(\alpha_3^2)^2)y_3^2-((\alpha_4^1)^2+(\alpha_4^2)^2)y_4^2 =$$
$$\frac{1}{2}((\alpha_3^1)^2+(\alpha_4^1)^2+(\alpha_3^2)^2+(\alpha_4^2)^2) = \frac{1}{2}(\|Q\alpha^1\|^2+\|Q\alpha^2\|^2)\,.$$

□

Lemma 21. *If in addition to assumptions of Lemma 20 we have*

$$(\alpha_1^1)^2+(\alpha_2^1)^2+(\alpha_1^2)^2+(\alpha_2^2)^2 \ge \frac{p}{2(p-1)}((\alpha_3^1)^2+(\alpha_4^1)^2+(\alpha_3^2)^2+(\alpha_4^2)^2)\,, \tag{6.17}$$

then

$$\langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^1,\alpha^1\rangle + \langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^2,\alpha^2\rangle \le 0\,. \tag{6.18}$$

Proof. Denote $H = (\alpha_1^1)^2+(\alpha_2^1)^2+(\alpha_1^2)^2+(\alpha_2^2)^2$, $K = (\alpha_3^1)^2+(\alpha_4^1)^2+(\alpha_3^2)^2+(\alpha_4^2)^2)$. Then Lemma 20 claims that

$$\langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^1,\alpha^1\rangle + \langle\frac{d^2\Phi(a)}{d^2\,a}\alpha^2,\alpha^2\rangle$$
$$\le -p(p-1)(H-K) - \frac{1}{2}p(p-2)K = -p(p-1)(H-\frac{p}{2(p-1)}K)\,.$$

We are done. □

Take the expectation of (6.10). We get then

$$\mathcal{E}(|Y_T|^p-(p-1)^p|X_T|^p)=\Phi(\Re X_0,\Im X_0,\Re Y_0,\Im Y_0)+\frac{1}{2}\int_0^T\frac{d^2\Phi}{d\,a^2}\overrightarrow{\alpha_k}\cdot\overrightarrow{\alpha_l}\,dt\,. \quad (6.19)$$

or

$$\mathcal{E}(|Y_T|^p-(p-1)^p|X_T|^p)=$$
$$\Phi(\Re X_0,\Im X_0,\Re Y_0,\Im Y_0)+\frac{1}{2}\int_0^T\langle\frac{d^2\Phi}{d\,a^2}\alpha^1,\alpha^1\rangle\,dt+\frac{1}{2}\int_0^T\langle\frac{d^2\Phi}{d\,a^2}\alpha^2,\alpha^2\rangle\,dt\,. \quad (6.20)$$

Choosing $\tau\,:\,4\,\tau^2\,\frac{p}{2(p-1)}=1$, that is $\tau=\sqrt{\frac{p-1}{2p}}$, and examining (6.13)–(6.15), we see that α^1,α^2 satisfy the assumptions of Lemmata 20, 21. Therefore the integrand is non-positive and, taking into consideration that $\Phi(\Re X_0,\Im X_0,\Re Y_0,\Im Y_0)=\Phi(\Re U_f(0,T),\Im U_f(0,T),...)\to 0$, $T\to\infty$, we get

$$\left(\sqrt{\frac{p-1}{2p}}\right)^p\mathcal{E}|A\star U_f(W_T,0)|^p\le(p-1)^p\mathcal{E}|U_f(W_T,0)|^p+o(1)\,.$$

Combining this with (6.9) we get the following theorem belonging to Banuelos and Janakiraman [BaJa1]:

Theorem 22.

$$\|\mathcal{B}f\|_p\le\sqrt{2p(p-1)}\|f\|_p\,.$$

7. Introducing the Bellster.

Here is a function $B_0=B_0(x_1,x_2,x_3,x_4)$ defined in $\Omega_0:=\{(x_1,x_2,x_3,x_4)\in\mathbb{R}^4\,:\,|x_2|^p\le x_1,|x_4|^q<x_3\}$.

$$B_0=\frac{(p-1)+(h-p)y_2^qh^{q-1}+y_1y_2[(p-1)(y_2h)^q+(1-ph]}{|1-(y_2h)^q|}x_1^{1/p}x_3^{1/q},$$

where $y_1=|x_2|/x_1^{1/p}$, $y_2=|x_4|/x_3^{1/q}$, and h - solves the following equation:

$$(p-1)y_2^{2q-1}h^{2q-2}+y_1y_2^qh^q-p(1+y_1y_2)y_2^{q-1}h^{q-1}+y_2^{q-1}h^{q-2}+(p-1)y_1=0.$$

Find y_1 and plug into the formula for B_0, then one gets another form of B_0 (with another denominator!):

$$B_0 = \mathrm{sign}(1-y_2h)\frac{(1-h)^2y_2^qh^{q-2}-(p-1)^2(1-y_2^qh^{q-1})^2}{py_2^qh^{q-1}-y_2^qh^q-(p-1)}x_1^{1/p}x_3^{1/q}.$$

Let us consider the following map of $\mathbb{R}^6$ into $\mathbb{R}^4$:

$$s(a_1,a_2,a_3,a_4,a_5,a_6) = (a_1, \sqrt{a_2^2+a_3^2}, a_4, \sqrt{a_5^2+a_6^2})\,.$$

Let

$$B := B_0 \circ s\,.$$

Its domain of definition is

$$\Omega = s^{-1}(\Omega_0) = \{a \in \mathbb{R}^6 : (a_2^2+a_3^2)^{\frac{p}{2}} \le a_1\,, (a_5^2+a_6^2)^{\frac{q}{2}} \le a_4\}\,.$$

Let d^2B_0, d^2B denote the Hessians of B_0, B correspondingly. These are 4×4 and 6×6 matrices.

Theorem 23. *Let α be an arbitrary vector in $\mathbb{R}^4$, $\alpha = (\alpha_1,\alpha_2,\alpha_3,\alpha_4)$, then*

$$-(d^2B_0\,\alpha,\alpha) \ge 2\,|\alpha_2||\alpha_4|\,,$$

moreover, at every point $x \in \Omega_0$ one of the quadratic forms

$$-(d^2B_0\,\alpha,\alpha) \pm 2\,\alpha_2\cdot\alpha_4$$

becomes "saturated", namely, there exists a non-zero vector α, where the last expression vanishes.

From this theorem it is easy to see the following

Theorem 24. *Let α be an arbitrary vector in $\mathbb{R}^6$, $\alpha = (\alpha_1,\alpha_2,\alpha_3,\dots,\alpha_6)$, then*

$$-(d^2B\,\alpha,\alpha) \ge 2\,|(\alpha_2,\alpha_3)||(\alpha_5,\alpha_6)|\,,$$

where $|(,)|$ denotes the Euclidean norm of the vector $(,)$.

We will call function B the "Bellster".

8. Questions about the Bellster

Notice that $p \geq 2$ and Bellster depends on p.

1. Let us perform the following operation on 6×6 matrix d^2B. Consider the entries $(d^2B)_{23}, (d^2B)_{32}$ and make them zero. New matrix will be called N.

Question 1. Is it true that

$$-(N\alpha,\alpha) \geq 2R\,|(\alpha_2,\alpha_3)||(\alpha_5,\alpha_6)|$$

with $R = R_p > 1$? What I mean is the following, may be R become bigger than 1 when $p\to\infty$? May be $\lim_{p\to\infty} R_p = 2$? or at least $\sqrt{2}$? Or may be this R_p is **always** close to $\sqrt{2}$ or even to 2 when $p > 2$? The last suggestion is really too good to be true.

2. Related question. Take **two** 6-vectors: $\alpha^1 = (\alpha^1_1,\dots,\alpha^1_6)\,, \alpha^2 = (\alpha^2_1,\dots,\alpha^2_6)$. Here is a very outrageous

Question 2. Is it true that

$$-(d^2B\,\alpha^1,\alpha^1) - (d^2B\,\alpha^2,\alpha^2) \geq$$

$$2\,[(\alpha^1_2)^2+(\alpha^1_3)^2+(\alpha^2_2)^2+(\alpha^2_3)^2+2(\alpha^1_2\alpha^2_3-\alpha^1_3\alpha^2_2)]^{\frac12}\,[(\alpha^1_5)^2+(\alpha^1_6)^2+(\alpha^2_5)^2+(\alpha^2_6)^2+2(\alpha^1_5\alpha^2_6-\alpha^1_6\alpha^2_5)]^{\frac12}\ ?$$

2a. May be the previous inequality holds when p becomes large?

3. Below we discuss the last inequality and see that may be it is too good to be true — our Bellster is not still enough of a monster probably to have it in such a strong form. But we have

Question 3. Does there exists a **super-bellster** $\mathcal{B}$ such that

$$-(d^2\mathcal{B}\,\alpha^1,\alpha^1) - (d^2\mathcal{B}\alpha^2,\alpha^2) \geq$$

$$2\,[(\alpha^1_2)^2+(\alpha^1_3)^2+(\alpha^2_2)^2+(\alpha^2_3)^2+2(\alpha^1_2\alpha^2_3-\alpha^1_3\alpha^2_2)]^{\frac12}\,[(\alpha^1_5)^2+(\alpha^1_6)^2+(\alpha^2_5)^2+(\alpha^2_6)^2+2(\alpha^1_5\alpha^2_6-\alpha^1_6\alpha^2_5)]^{\frac12}\ ?$$

The positive answer to this problem solves $p-1$ problem.

4. Is there a simple reason to see that super-bellster cannot exist?

9. Discussion of questions

Consider question 2 and plug vectors such that $\alpha^1 = \alpha^2 =: \alpha$. Then the brackets $(\alpha^1_2\alpha^2_3 - \alpha^1_3\alpha^2_2))$ disappear and we come to the inequality

$$-2(d^2B\,\alpha, \alpha) \geq 2\cdot 2[(\alpha_1)^2 + (\alpha_2)^2]^{\frac{1}{2}}[(\alpha_3)^2 + (\alpha_4)^2]^{\frac{1}{2}}\,,$$

which is correct by Theorem 21.3.

On the other hand, let us plug vectors α^1, α^2 such that

$$\alpha^1_2 = \alpha^2_3\,, \alpha^1_3 = -\alpha^2_2\,.$$

And all other α^1's equal to all other α^2's. Call $\alpha = \alpha^1$. Then in the left hand side we get exactly $-2(N\alpha, \alpha)$. In the right hand side we get $2\cdot 2\cdot\sqrt{2}[\alpha_2^2+\alpha_3^2]^{\frac{1}{2}}[\alpha_5^2+\alpha_6^2]^{\frac{1}{2}}$, and we come to question 1 with $R \geq \sqrt{2}$ for all $p \geq 2$. It would be extremely good. But if this is false, we still can hope for 2a) above.

What I mean by that is that inequality in question 2 can become correct when p becomes large.

If our bellster is not good enough for this, one should try to build a super-bellster of question 3.

10. Martingale transforms and Burkholder functions.

We qoute Burkholder: "It must have been known to Alexander Calder that it is possible to design a mobile that can be hung initially in a small room but which, if it is to move freely through all of its possible configurations, will have to be hung anew in an exceedingly large room. There is a close mathematical analogue. To each possible configuration of a mobile made with strings, rods, and weights, there corresponds a martingale with a similar arrangement of successive centers of gravity and this martingale is the martingale transform of the martingale corresponding to the initial configuration. It is easy to see, either by looking first at mobiles or directly to martingales, that there do exist small martingales with large transforms".

Burkholder found the exact relations between the sizes of martingales and their transforms.

Let r_k denote Rademacher functions on $[0, 1)$, that is $r_k(t) = \pm$ based on the decomposition of t in base 2, if the k-th figure is zero, $r_k(t) = 1$, otherwise, if it is one, $r_k(t) = -1$. Notice that r_k is orthogonal to $\phi(r_1, ..., r_{k-1})$, for all, say, bounded ϕ.

Consider the functions $\Phi := \sum_k r_k\, d_k(r_1, ..., r_{k-1})$, where sum is finite. Let α_k be numbers ± 1. Our first martingale transform is $\sum_k \alpha_k\, r_k\, d_k(r_1, ..., r_{k-1})$.

Here is a celebrated theorem of Burkholder

Theorem 25. *Let $1 < p < \infty$.*

$$\sup_{\alpha_k = \pm 1} \| \sum_k \alpha_k\, r_k\, d_k(r_1, ..., r_{k-1})\|_p \le (p^* - 1) \| \sum_k r_k\, d_k(r_1, ..., r_{k-1})\|_p\,.$$

Here $p^* := \max(p, \frac{p}{p-1})$. Constant $p^* - 1$ is sharp for every p.

In [Bu1] the statement is different, and this is why we want to explain the complete similarity of statements.

Consider the sequence $\{t_n\}$ of real bounded functions, each of which depends only on finitely many Rademacher functions. Let $t_1, ..., t_{k-1}$ are functions of $r_1, ..., r_{n_k}$ Let t_k be orthogonal to $\phi(r_1, ..., r_{n_k})$ for every bounded ϕ. Then we call $\{t_n\}$ the martingale difference sequence and sequence $f = \{f_n\}$, $f_n = \sum_{k=1}^n t_k$, $n = 1, 2, ...$ is called martingale. (Notice that $t_k = r_k d(r_1, ..., r_{k-1})$ above is a particular case of the martingale difference sequence.) Let $v_k = v_k(t_1, ..., t_{k-1})$ be a sequence of bounded functions. The sequnece $g = \{g_n\}$, $g_n = \sum_{k=1}^n v_k\, t_k$ is called the martingale transform of f by v. This is the main result of [Bu1], [Bu7]:

Theorem 26. *Let $1 < p < \infty$. For every real v such that $|v_k| \le 1$ and every n one has*

$$\| \sum_{k=1}^n v_k\, t_k\|_p \le (p^* - 1) \| \sum_{k=1}^n t_k\|_p\,.$$

Constant $p^* - 1$ is sharp for every p.

We want to show that Theorem 19 implies Theorem 26. The converse is just obvoius because the sequence $r_n d(r_1, ..., r_{n-1})$ is a particular case of t_n and because transforming by *constant* functions $v_k = \alpha_k$ is a particular case of transforming by *general* functions $v_k, |v_k| \le 1$. However, it turns out that it is all the same.

1. A trivial remark is that of course functions v_k can be considered taking only two values $\pm$. In fact, given $x \in [-1, 1]$, we write $x = \sum_{j=1}^\infty 2^{-j} c_j(x)$, where

$c_j : [-1,1] \to \{-1,1\}$. Let $v_k^j(t) := c_j(v_k(t))$. For each j consider martingale transforms of f by v_k^j, and call it g^j. If we have

$$\|g_n^j\|_p \le (p^* - 1)\|f_n\|_p$$

we immeditely conclude that

$$\|g_n\|_p \le (p^* - 1)\|f_n\|_p$$

because $g_n = \sum_{j=1}^{\infty} 2^{-j} g_n^j$ for every n.

2. We reduced v_k to the case of functions assuming two values $\pm$. We want to make them constant functions: either 1 or -1. Consider the new martingale difference sequence $T_{2n-1} := (1+v_n)t_n/2$, $T_{2n} := (1-v_n)t_n/2$. To see that it is a martingale difference sequence we need to check that T_{2n-1} is orthogonal to $\phi(T_1, ..., T_{2n-2})$ and T_{2n} is orthogonal to $\phi(T_1, ..., T_{2n-2}, T_{2n-1})$. Let us check the latter, for example. Looking at

$$\int (1-v_n)t_n\, \phi(T_1, ..., T_{2n-2}, (1+v_n)t_n/2)\, dt$$

we see that it is $\int t_n\,(1 - v_n(t_1, ..., t_{n-1})\Phi(t_1, .., t_{n-1}) = 0$. When $v_n = 1$ both integrand vanish (and so they are equal), and when $v_n = -1$, $(1+v_n)t_n/2 = 0$ and $\phi(T_1, ..., T_{2n-2}, (1+v_n)t_n/2) = \phi(T_1, ..., T_{2n-2}, 0) =: \Phi(t_1, ..., t_{n-1})$.

Now $F_{2n} = f_n$ obviously. Let us transform martingale $\{F_n\}$ by sequence of constant functions $\{1, -1, 1, -1, ...\}$. We get $\{G_n\}, G_{2n} = \sum_{k=1}^{2n}(-1)^{k-1}T_k = \sum_{k=1}^{n} v_k t_k = g_n$. Threfore, if we have

$$\|G_n\|_p \le (p^* - 1)\|F_n\|_p$$

then the same is true for g_n, f_n. So we reduced Theorem 26 to the transforms by constant functions each being either 1 or -1. Moreover, we need only one special transform: by $1, -1, 1, -1, ...$!

3. So our functions v_k are now constant functions $v_k = 1$ or -1. The last thing we should understand is why arbitrary martingale diffrence sequence $\{t_k\}$ can be reduced to special sequences $\{r_k\, d(r_1, ..., r_{k-1})\}$. For this we just need an obvious approximation remark. Let $t_1, ..., t_{k-1}$ are functions of $r_1, ..., r_{n_k}$. Any function t orthogonal to any $\phi(r_1, .., r_{n_k})$ can be approximated as well as we like by $\psi(r_{n_k+1}, ..., r_{n_{k+1}})$ if n_{k+1} is chosen to be large enough. This is a consequence of the density of Rademacher functions in L^2. Now write $\psi(r_{n_k+1}, ..., r_{n_{k+1}}) =$

$\sum_{m=n_k+1}^{n_{k+1}} r_m\, d_m(r_1,...,r_{m-1})$. This is always possible (look at Haar decomposition of ψ for example). Now the estimate

$$\|\sum_{k=1}^{K}\sum_{m=n_k+1}^{n_{k+1}} v_k\, r_m\, d_m(r_1,...,r_{m-1})\|_p \le (p^*-1)\|\sum_{k=1}^{K}\sum_{m=n_k+1}^{n_{k+1}} r_m\, d_m(r_1,...,r_{m-1})\|_p$$

becomes

$$\|\sum_{k=1}^{K} v_k\, t_k\| \le (p^*-1)\|\sum_{k=1}^{K} t_k\|_p + \varepsilon$$

with arbitrary $\varepsilon > 0$ (which reflects the descrepancy coming from the approximation). Notice that the sequence of numbers $\alpha_m = v_k$, $m \in [n_k+1, n_{k+1}]$ is no more a simple unique sequence $\{1,-1,1,-1,...\}$.

We proved that Theorem 19 implies Theorem 26. In other words the supremum of norm of transforms in Theorem 26 is equal to the supremum of norm of transforms in Theorem 19.

11. The estimate from below on multipliers of the form $\frac{(A\xi,\xi)}{|\xi|^2}$.

In this section we finish **computing** the norm of several multipliers in L^p. Let us recall the reader that multipliers whose norm can be computed are very rare. We know the celebrated Pichorides [Pi] theorems, and a result of Kalton–Verbitsky [KV]. May be this is all, apart from those found by the combination of works of Dragicevic, Nazarov, Volberg, Banuelos–Hernandez, Geiss–Montgomery-Smith–Saksman[NV], [DV1], [DV2], [**?**] and [GMSS] and treated in this section.

Operator $\sum_{k,l=1}^{d} a_{kl} R_k R_l$ is the multiplier operator

$$M_m f = (m(\xi)\,\hat{f}(\xi))\check{}$$

with $m(\xi) = \frac{(A\xi,\xi)}{|\xi|^2}$ and matrix $A = (a_{kl})_{k,l=1}^{d}$. These multipliers m are real analytic on the sphere S^{d-1}, the only point of discontinuity of function m in $\mathbb{R}^d$ is the origin, and they are homogeneous of degree 0:

$$m(t\xi) = m(\xi)\,,\ t \neq 0\,.$$

Thre are two important and simple facts that we have to use about multipliers.

Lemma 27. *Let $L : \mathbb{R}^d \to \mathbb{R}^d$ be a linear transformation. Then $\|M_m\|_p = \|M_{m\circ L}\|_p$.*

Proof. By applying the simple identity

$$M_{m\circ L}f(x) = (M_m(f\circ L^T))((L^T)^{-1}x)$$

we immediately finish the proof. □

All functions on the torus $\mathbb{T}^d$ will be assumed to have zero average, and $L^p_0(\mathbb{T}^d)$ denotes the functions from $L^p(\mathbb{T}^d)$ with zero average. The second fact deals with multipliers on $L^p(\mathbb{T}^d)$ versus multipliers on $L^p(\mathbb{R}^d)$. Given m and any trigonometric polynomial f on $\mathbb{T}^d$ with zero avaerage, we define

$$T_m : L^p_0(\mathbb{T}^d) \to L^p(\mathbb{T}^d)$$

by

$$T_m f = \sum_{n\in\mathbb{Z}^d} m(n)\hat{f}(n)e^{i(n,\theta)}\,,$$

$\theta = (\theta_1, ..., \theta_d) \in [0, 2\pi)^d$.

Lemma 28. $\|T_m\|_p = \|M_m\|_p$

Proof. Let us check $\le$ first. Given a trigonometric polynomials $F(\theta), G(\theta)$ we consider $f(x) := F(x)e^{-\pi\varepsilon/2|x|^2}$, $g(x) := G(x)e^{-\pi\varepsilon/2|x|^2}$, where $F(x), G(x)$ are periodic extensions of F, G onto $\mathbb{R}^d$. Then the following holds

$$\int_{[0,2\pi)^d} F(\theta)\overline{G(\theta)}\,d\theta = \lim_{\varepsilon\to 0} \varepsilon^{\frac{d}{2}} \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx\,. \tag{11.1}$$

It is sufficient to prove this for $F = e^{i(m,x)}, G = e^{i(k,x)}$. Then LHS $= \delta_{km}$. On the other hand

$$\varepsilon^{\frac{d}{2}} \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx = \varepsilon^{\frac{d}{2}} \int \frac{1}{(\varepsilon/4)^d} e^{-\pi\frac{2|\xi-m|^2}{\varepsilon}} e^{-\pi\frac{2|\xi-k|^2}{\varepsilon}}\,,$$

$= (\frac{4}{\varepsilon})^{\frac{d}{2}} \int e^{-\pi\frac{4|\xi-m|^2}{\varepsilon}}$ if $m = k$. The last expression tends to 1 when $\varepsilon \to 0$. If $m \ne k$ the expression in display formula obviously goes to zero.

Having proved (11.1) we are done. In fact, then $(T_mF, G) = (M_mf, g)$ for any two given trigonometric polynomials F, G. We are left to see that (and the same for G)

$$\|F\|^p_{L^p(\mathbb{T}^d)} = \lim_{\varepsilon\to 0} \varepsilon^{\frac{d}{2}} \int_{\mathbb{R}^d} |F(x)|^p e^{-\varepsilon\pi|x|^2}\,dx\,.$$

This follows from

$$\int_{[0,2\pi)^d} g(x)\,dx = \lim_{\varepsilon\to 0} \varepsilon^{\frac{d}{2}} \int_{\mathbb{R}^d} g(x)e^{-\varepsilon\pi|x|^2}\,dx\,,$$

where $g(x)$ in the RHS are periodic extensions of g. To prove the latter realations let us write

$$\varepsilon^{\frac{d}{2}}\int_{\mathbb{R}^d} g(x)e^{-\varepsilon\pi|x|^2}\,dx = \int_{[0,2\pi)^d} g(x)\sum_{n\in\mathbb{Z}^d}\varepsilon^{\frac{d}{2}}e^{-\varepsilon\pi|x-n|^2}\,dx =$$

$$\int_{[0,2\pi)^d} g(x)\sum_{n\in\mathbb{Z}^d} e^{-\frac{\pi|n|^2}{\varepsilon}}e^{i(n,x)} = \int_{[0,2\pi)^d} g(x)\,dx + o(1)\,,$$

as $\sum_{n\in\mathbb{Z}^d\setminus\{0\}} e^{-\frac{\pi|n|^2}{\varepsilon}} = o(1)$.

□

We used in the proof the **Poisson summation formula**:

$$\sum_{n\in\mathbb{Z}^d}\hat f(n)e^{i(n,x)} = \sum_{n\in\mathbb{Z}^d} f(x-n)$$

for $f = e^{-\varepsilon\pi|x|^2}$ valid for every Schwarz function f.

Let now

$$m^+ := \max_{\xi\in\mathbb{R}^d} m(\xi)\,,\ m^- = \min_{\xi\in\mathbb{R}^d} m(\xi)\,,$$

$$v^+, v^- \text{ vectors where this maximum and minimum are attained}\,,$$

$$c = \frac{m^+ + m^-}{2}\,,\ \ell := \frac{m^+ - m^-}{2}\,,\ I_m = [c-\ell, c+\ell] = [m^-, m^+]\,.$$

Consider martingale transforms as in Theorem 19 but with numbers α_k not in $[-1,1]$ but in $[c-\ell, c+\ell] = [m^-, m^+]$. In other words, introduce (all sums are finite) operators

$$MT_\alpha\Big(\sum r_k\, d(r_1,...,r_{k-1}\Big) = \sum \alpha_k\, r_k\, d(r_1,...,r_{k-1})$$

where each α_k is a number in I_m.

It would be very interesting to find

$$B(a,b,p) := \sup_{\alpha_k\in[a,b]} \|MT_\alpha\|_p = \ell(p^*-1)\,.$$

But for some intervals we know it! We know from Theorem 19 that the following holds

Theorem 29. *If* $c=0$ *then*

$$B(-\ell,\ell,p) = \ell(p^*-1)\,.$$

Here is a remarkable result of Geiss–Montgomery-Smith–Saksman, which sometimes allows us to **compute** the norm of the multiplier $\|M_m\|_p$.

Theorem 30. $\|M_m\|_p \geq \sup_{\alpha_k \in I_m} \|MT_\alpha\|_p$.

Example. This theorem and our Theorem 6 prove

$$\|(R_1^2 - R_2^2)\cos\theta + 2R_1R_2\sin\theta\|_p = p^* - 1$$

for all θ.

In fact, for $m(\xi) = ((\xi_1)^2 - (\xi_2)^2)\cos\theta + 2\xi_1\xi_2\sin\theta)(\xi_1^2 + \xi_2^2) = (A_\theta\xi, \xi)/|\xi|^2$ we have $I_m = [-1, 1]$. We just notice that matrix A_θ is a flip followed by rotation, so it is unitary and there are vectors v^+, v^- such that

$$A_\theta v^+ = v^+,\ A_\theta v^- = -v^- .$$

11.1. **The proof of Theorem 30.** The proof will follow an idea of Bourgain [B1] of augmanting the number of variables.

By Lemma 28 it is enough to estimate

$$\|T_m\|_P \geq \sup_{\alpha_k \in I_m} \|MT_\alpha\|_p . \tag{11.2}$$

First we need Bourgain's lemma. Let $Q := \mathbb{T}^d$. Consider the set $\mathcal{T}_k$ of trigonometric polynomilas on Q^k of the following form

$$\Phi(\theta_1, ..., \theta_k) = \sum_{l \in \mathbb{Z}^d,\, l \neq 0} \Phi_l(\theta_1, ..., \theta_{k-1}) e^{i(l,\theta_k)} ,$$

where $\Phi_l(\theta_1, ..., \theta_{k-1})$ are arbitrary trigonometric polynomials on Q^{k-1}.

Operator T_m can be naturally extended from usuual trigonometric polynomials on Q without the free term to this new family $\mathcal{T}_k$ of trigonometric polynomials on Q without the free term with coefficients in trigonometric polynomials on Q^{k-1}. We call this extension B_m:

$$B_m\Phi(\theta_1, ..., \theta_k) = \sum_{l \in \mathbb{Z}^d,\, l \neq 0} \Phi_l(\theta_1, ..., \theta_{k-1}) m(l) e^{i(l,\theta_k)} = \sum_{l \in \mathbb{Z}^d,\, l \neq 0} \Phi_l(\theta_1, ..., \theta_{k-1}) T_m(e^{i(l,\theta_k)}) .$$

Now consider *any* trigonometric polynomial on Q^k:

$$\Psi(\theta_1, ..., \theta_k) = \sum_{(l_1,..,l_k) \in (\mathbb{Z}^d)^k} x_{l_1,...,l_k} e^{i(l_1,\theta_1)} e^{i(l_2,\theta_2)} ... e^{i(l_k,\theta_k)}$$

such that it has zero average. Then obviously

$$\Psi = \Phi_k(\theta_1, ..., \theta_k) + \Phi_{k-1}(\theta_1, ..., \theta_{k-1}) + ... + \Phi_1(\theta_1) ,$$

where $\Phi_k \in \mathcal{T}_k$. We can write this representation clearly in the unique way. Then B_m extends to the class $\mathcal{P}_k$ of all trigonometric polynomials on Q^k with zero avaerage by

$$B_m\Psi = B_m\Phi_k(\theta_1, ..., \theta_k) + B_m\Phi_{k-1}(\theta_1, ..., \theta_{k-1}) + ... + B_m\Phi_1(\theta_1)\,.$$

As usual we call

$$\|B_m\|_p = \sup_{\Phi\in\mathcal{P}_k, \Phi\neq 0} \frac{\|B_m\Phi\|_p}{\|\Phi\|_p}$$

the norm of B_m. Obviously it is at least as large as $\|T_m\|_p$ as B_m extends T_m. Bourgain's lemma claims that the norm stays the same.

Lemma 31. *Let m be continuous on S^{d-1}. Then $\|B_m\|_p = \|T_m\|_p$.*

Proof. Let N be a very large integer. We will tend it to infinity later. Let $\eta \in Q = [0, 2\pi)^d$ be one extra variable, and let $f \in \mathcal{P}_K$. Consider one parametric family of trigonometric polynomials $f_\eta(\theta_1, ..., \theta_k) = f(\theta_1 + N\eta, ..., \theta_k + N^k\eta)$. Of course

$$\frac{1}{(2\pi)^d}\int_{[0,2\pi)^d} d\eta \|f_\eta\|^p_{L^p(Q^k)} =$$

$$\|f_\eta\|^p_{L^p(Q^k)}\,, \quad \frac{1}{2\pi}\int_0^{2\pi} d\eta \|B_m\, f_\eta\|^p_{L^p(Q^k)} =$$

$$\|B_m\, f_\eta\|^p_{L^p(Q^k)}\,.$$

This is just Fubini's theorem. Therefore, to compare $\|B_m\, f\|^p_{L^p(Q^k)}$ with $\|f\|^p_{L^p(Q^k)}$ is the same as to compare these averages. Let us compare each individual $\|B_m\, f_\eta\|^p_{L^p(Q^k)}$ with $\|f_\eta\|^p_{L^p(Q^k)}$. Let us consider the trigonometric polynomial of "one" variable η given by $(\theta := (\theta_1, ..., \theta_k))$

$$F_\theta(\eta) = f_\eta(\theta_1, ..., \theta_k)\,.$$

Let us compare $T_mF_\theta(\eta)$ and $B_m\, f_\eta(\theta_1, ..., \theta_k)$.

$$B_m\, f_\eta(\theta_1, ..., \theta_k) = \sum_{k=1}^{K} \sum_{(l_1,..,l_k)\in(\mathbb{Z}^d)^k} x_{l_1,...,l_k} m(l_k) e^{i(l_1,\theta_1+N\eta)} e^{i(l_2,\theta_2+N^2\eta)} ... e^{i(l_k,\theta_k+N^k\eta)}\,.$$

$$T_m\, f_\theta(\eta) =$$

$$\sum_{k=1}^{K} \sum_{(l_1,..,l_k)\in(\mathbb{Z}^d)^k} x_{l_1,...,l_k} m(N\, l_1 + N^2\, l_2 + ... + N^k\, l_k) e^{i(l_1,\theta_1+N\eta)} e^{i(l_2,\theta_2+N^2\eta)} ... e^{i(l_k,\theta_k+N^k\eta)}\,.$$

Notice that 0-homogenuity of m implies

$$m(N\, l_1 + N^2\, l_2 + ... + N^k\, l_k) = m(l_k + \frac{1}{N} l_{k-1} + ... + \frac{1}{N^{k-1}} l_1) \to m(l_k) \text{ when } N \to \infty\,.$$

In particular, given f for extremely large N we have that $B_m\, f_\eta(\theta_1, ..., \theta_k)$ is uniformly close to $T_m\, f_\theta(\eta)$ as much as we wish. So $\int_Q d\eta \|B_m\, f_\eta\|^p_{L^p(Q^k)}$ is as close to $\int_{Q^k} \|T_m f_\theta(\eta)\|^p_{L^p(Q)}$. Hence

$$\|B_m\, f(\theta_1, ..., \theta_k)\|^p_{L^p(Q^k)} = \int_Q d\eta \|B_m\, f_\eta\|^p_{L^p(Q^k)} \leq \int_{Q^k} \|T_m f_\theta(\eta)\|^p_{L^p(Q)} + \varepsilon$$

$$\leq ||T_m\|^p_p \int_{Q^k} \|f_\theta(\eta)\|^p_{L^p(Q)} + \varepsilon = ||T_m\|^p_p \|f(\theta_1, ..., \theta_k\|^p_{L^p(Q)} + \varepsilon\,.$$

We already said that the last and the first equalitis are just Fubini's theorem. Making $N \to \infty$ we make $\varepsilon \to 0$ and Bourgain's lemma is proved.

□

Now we are ready to prove that for $m \in C(S^{d-1})$ we have

$$\|T_m\|_p \geq \sup_{\alpha_k \in I_m} \|MT_\alpha\|_p\,.$$

By Lemma 31 we need to show that

$$\|B_m\|_p \geq \sup_{\alpha_k \in I_m} \|MT_\alpha\|_p\,. \tag{11.3}$$

The nice thing about B_m is that it does not depend on the number of variables. Let now

$$m^+ := \max_{\xi \in S^{d-1}} m(\xi)\,,\, m^- = \min_{\xi \in S^{d-1}} m(\xi)\,,$$

v^+, v^- unit vectors where this maximum and minimum are attained.

So let us fix K, $\alpha_1, ..., \alpha_K$, and functions $d_1(\varepsilon_1), ..., d_K(\varepsilon_1, ..., \varepsilon_K)$ such that the supremum $S := \sup_{\alpha_k \in I_m} \|MT_\alpha\|_p$ is almost achieved on $\sum_k \varepsilon_k d_{k-1}(\varepsilon_1, .., \varepsilon_{k-1})$. Namely,

$$\|\sum_{k=1}^{K+1} m^{\alpha_k}\, \varepsilon_k d_{k-1}(\varepsilon_1, .., \varepsilon_{k-1})\| \geq S \|\sum_{k=1}^{K+1} \varepsilon_k d_{k-1}(\varepsilon_1, .., \varepsilon_{k-1})\|_p - \delta\,. \tag{11.4}$$

Choose $a = (n_1, ..., n_d) \in \mathbb{Z}^d$, $b = (n'_1, ..., n'_d) \in \mathbb{Z}^d$ such that

$$|m(a) - m^+| \leq \varepsilon\,,\, |m(b) - m^-)| \leq \varepsilon\,. \tag{11.5}$$

Consider functions $\mathrm{sgn}(\theta) = \sum_{n \in \mathbb{Z} \setminus \{0\}} p_n e^{in\theta}$ and put

$$\psi_+(\theta) = \sum_{n \in \mathbb{Z} \setminus \{0\}} p_n e^{i(na, \theta)}\,, \psi_-(\theta) = \sum_{n \in \mathbb{Z} \setminus \{0\}} p_n e^{i(nb, \theta)}\,,\ \theta \in Q = \mathbb{T}^d\,.$$

Consider the sequence $\alpha_1, \alpha_2, ..., \alpha_k, ..., \alpha_{K+1}$ as a sequence of $+$ and $-$ (recall that each α_k is $\pm$) and correspond to it the sequence of functions on $Q = \mathbb{T}^d$ $\psi_{\alpha_1}(\theta), ..., \psi_{\alpha_k}(\theta), ..., \psi_{\alpha_{K+1}}(\theta)$. Call it for brevity $\psi_1(\theta), ..., \psi_k(\theta), ..., \psi_{K+1}(\theta)$. These are not trigonometric polynomials, but we will approximate these functions by trigonometric polynomilas in the future.

Now let us just notice that the joint distribution of

$$\psi_1(\theta_1), ..., \psi_k(\theta_k), ..., \psi_{K+1}(\theta_{K+1})$$

on Q^{K+1} is the same as for Bernoulli random variables $\varepsilon_1, ..., \varepsilon_{K+1}$.

Imitating (11.4) we consider the function

$$f(\theta_1, ..., \theta_{K+1}) := \sum_{k=1}^{K+1} \psi_k(\theta_k) d_{k-1}(\psi_1(\theta_1), .., \varepsilon_{k-1}(\theta_{k-1})\,.$$

Our remark about the joint distribution shows that

$$\|f(\theta_1, ..., \theta_{K+1})\|_{L^p(Q^{K+1})} = \|\sum_{k=1}^{K+1} \varepsilon_k d_{k-1}(\varepsilon_1, .., \varepsilon_{k-1})\|_p\,. \tag{11.6}$$

Consider now

$$B_m f(\theta_1, ..., \theta_{K+1}) := \sum_{k=1}^{K+1} B_m(\psi_k(\theta_k) d_{k-1}(\psi_1(\theta_1), .., \varepsilon_{k-1}(\theta_{k-1}))\,.$$

Of course, B_m, T_m were defined only on trigonometric polynomials with zero avaerage, but being a bounded operator they can be extended to their closure in $L^p_0(Q^{K+1}), L^p(Q)_0$ correspondingly. And functions f, ψ_k are of course in $L^p_0(Q^{K+1}), L^p(Q)_0$ correspondingly. Moreover, let us notice that

$$B_m(\psi_k(\theta_k) d_{k-1}(\psi_1(\theta_1), .., \varepsilon_{k-1}(\theta_{k-1})) = d_{k-1}(\psi_1(\theta_1), .., \varepsilon_{k-1}(\theta_{k-1}) T_m(\psi_k(\theta_k))$$

and

$$T_m(\psi_k(\theta_k)) = m(a)\psi_k(\theta_k) \text{ or } m(b)\psi_k(\theta_k)\,,$$

depending on whether $\alpha_k = 1$ or -1 (and correspondinfly $\psi_k(\cdot) = \mathrm{sgn}((a, \cdot))$ or $\psi_k(\cdot) = \mathrm{sgn}((b, \cdot))$). In fact, let $\alpha_k = 1$. Then

$$T_m\psi_k(\theta) = T_m(\sum_{n\in\mathbb{Z}\setminus\{0\}} p_n e^{i(na,\theta)}) = \sum_{n\in\mathbb{Z}\setminus\{0\}} m(na) p_n e^{i(na,\theta)} = m(a) \sum_{n\in\mathbb{Z}\setminus\{0\}} p_n e^{i(na,\theta)}$$

by homogenuity of m. The same when $\alpha_k = -1$, but with changing a to b.

So we can write (choosing ε in (11.5) as small as we wish) that

$$\|B_m(f(\theta_1, ..., \theta_{K+1}))\|_p =$$

$$\|B_m(\sum_{k=1}^{K+1}\psi_k(\theta_k)d_{k-1}(\psi_1(\theta_1),..,\varepsilon_{k-1}(\theta_{k-1}))\|_{L^p(Q^{K+1})} =$$
$$\|\sum_{k=1}^{K+1} m^{\alpha_k}\,\psi_k(\theta_k)\,d_{k-1}(\psi_1(\theta_1),..,\varepsilon_{k-1}(\theta_{k-1})\|_p - \delta\,,$$

where δ is arbitrarily small. By our remark about the distribution the right hand side here is equal to $\|\sum_{k=1}^{K+1} m^{\alpha_k}\,\varepsilon_k d_{k-1}(\varepsilon_1,..,\varepsilon_{k-1})\|$. Use now (11.4) and (11.6). We get

$$\|B_m(f(\theta_1,...,\theta_{K+1}))\|_p \geq S\|f(\theta_1,...,\theta_{K+1})\|_p - 2\,\delta$$

with arbitrary small δ.

Recall that $S = \sup_{\alpha_k\in I_m}\|MT_\alpha\|_p$ and, henceforth, we proved (11.3), which together with Lemma 31 finishes the proof of Theorem 30:

$$\|M_m\|_p \geq \sup_{\alpha_k\in I_m}\|MT_\alpha\|_p\,.$$

Remark. Of course by Lemma 27 we can think that $v^+ = e_1, v^- = e_2$, first two orts of $\mathbb{R}^d$. This simplifies the choice of functions ψ_+, ψ_-: $\psi_+ := \text{sgn}((e_1,\theta)), \psi_- := \text{sgn}((e_2,\theta))\,, \theta := (\theta_1,...,\theta_d)$.

12. Burkholder's Bellman function

We will show how to get Burkholder's function from [Bu1] by using Monge-Ampère equation. This method is quite different than the one in the series of Burkholder's papers [Bu1]–[Bu7], and as we will see, is quite a universal method.

13. Notations and definitions

Bellman function method in Harmonic Analysis was introduced by Burkholder for finding the norm in L^p of the Martingale transform. Later it became clear that the scope of the method is quite wide.

The technique, originated in Burkholder's papers [Bu1]–[Bu7], can be credited for helping to solve several old Harmonic Analysis problems and for unifying approach to many others. In the first category one would name the (sharp weighted) estimates of such classical operators as the Ahlfors–Beurling transform (Banuelos–Wang [BaWa1], Banuelos–Janakiraman [BaJa1], Banuelos–Mendez [BaMH], Nazarov–Volberg [NV1], Petermichl–Volberg [PV], Dragicevic–Volberg [DV2]) and the Hilbert and Riesz transforms (Petermichl [**?**], [**?**]). In the second category one can name all kind of dimension free estimates of weighted and unweighted Riesz transforms (see a vast literature in [DV1]–[DV3]). Roughly, Bellman function method makes apparent the hidden scaling properties of a given Harmonic Analysis problem. Conversely, given a Harmonic Analysis problem with certain scaling properties one can (formally) associate with is a non-linear PDE, the so-called Bellman equation of the problem.

Let us recall to the reader that in the series of papers [**?**], [Bu1]–[Bu7] Donald Burkholder investigated Martingale transform and gave the sharp bounds on this operator in various settings–but by similar methods. The methods were so novel and powerful that the influence of these articles will be felt for many years to come. The novelty was a key. One of the leading mathematician working in the domain of Harmonic Analysis told the second author that these papers of Burkholder "spin his head". In the book of Daniel Strook [Str] many pages are devoted to the technique developed by Burkholder in the abovementioned series of papers, and the reader can sense the same feeling. It is explained in [Str] that the simplest way to understand the sharp estimates of Martingale transform obtained by Burkholder is to operate with one of the so-called Burkholder's function:

$$u_p(x,y) = p(1-\frac{1}{p^*})^{p-1}(|y|-(p^*-1)|x|)(|x|+|y|)^{p-1}\,, \tag{13.1}$$

here $p^* := \max(p, \frac{p}{p-1}), 1 < p < \infty$.

However, the main question is of course how to get this function? Where did it come from? These questions are asked in [Str] as well. Of course, Burkholder explains in many details the way this function (and several of its relatives) are obtained. It is almost (but not quite) the least bi-concave majorant of function

$$|y|^{p-1} - (p^*-1)^p|x|^p\,. \tag{13.2}$$

It is obtained by solving a certain PDE and performing certain manipulations with the solution after that. The reader will find much more about u_p after reading this article, in particular in Sections 17, 17.1.

But it seems like the same questions persist even after this explanation. And a new question can appear: how wide is the applicability of the technique that Burkholder elaborated in [Bu1]–[Bu7]? There is a vague feeling that the area of applicability is quite wide. To make this feeling more precise one should look at the function above closer and see that it is a creature from another universe, which, initially, does not have too much in common with Harmonic Analysis. Burkholder function is a natural dweller of the area called Stochastic Optimal Control. It is a solution of a corresponding Bellman equation (or a dynamic programming equation) but in the setting, when the differential equations subject to control are not the usual ones. They are stochastic differential equations.

The reader can find some notes on this in [**?**], [VaVo], [VaVo3], [SlSt]. These notes explain why Stochastic Optimal Control is the right tool to work with a certain class of Harmonic Analysis problems. On the other hand, Stochastic Optimal Control problems generically can be reduced to solving a so-called Bellman PDE (and proving the so-called "verification theorems", but this is a second task). Bellman PDEs belong to the class of fully non-linear PDEs. Often they are PDEs of Monge–Ampère type. In the present article we would like to show the reader how to obtain Burkhloder functions (the one above and others from [Bu1]–[Bu7]) by reducing the search for them to solving certain Monge–Ampère equations. The scope of the application of the methods of Stochastic Optimal Control to Harmonic Analysis proved to be quite large. After Burkholder the first systematic application of this technique appeared in 1995 in the first preprint version of [**?**]. It was vastly developed in [NT] and in (now) numerous papers that followed. A small part of this literature can be found in the bibliography below.

We shall say that an interval I and a pair of positive numbers $\alpha^\pm$ ($\alpha^+ + \alpha^- = 1$) generate a pair of subintervals I^+ and I^- if $|I^\pm| = \alpha^\pm |I|$ and $I = I^- \cup I^+$. To every interval J and every sequence

$$\{\alpha_{n,m} \colon 0 < \alpha_{n,m} < 1, \quad 0 \le m < 2^n, \quad 0 < n < \infty\}$$

such that $\alpha_{n,2k} + \alpha_{n,2k+1} = 1$ we put in correspondence a family $\mathcal{I}$ of intervals $\{I_{n,m}\}$ such that any interval $I_{n,m}$ generates the pair $I_{n+1,2m} = I^-_{n,m}$, $I_{n+1,2m+1} = I^+_{n,m}$ with $\alpha^- = \alpha_{n+1,2m}$, $\alpha^+ = \alpha_{n+1,2m+1}$, starting with $I_{0,0} = J$. The symbol $\mathcal{I} = \mathcal{I}(\alpha)$ will denote the families of subintervals of J corresponding to a fixed choice of the numbers $\alpha_{n,m}$. For a special choice $\alpha_{n,m} = \frac{1}{2}$ we get the dyadic family $\mathcal{I} = \mathcal{D}$. Every family $\mathcal{I}$ has its own set of Haar functions:

$$\forall I \in \mathcal{I} \qquad h_{_I}(t) = \begin{cases} -\sqrt{\frac{\alpha^+}{\alpha^-|I|}} & \text{if } t \in I_-, \\ \sqrt{\frac{\alpha^-}{\alpha^+|I|}} & \text{if } t \in I_+. \end{cases}$$

If the family $\mathcal{I}$ is such that that the maximal length of the interval of n-th generation (i.e., $\max\{|I_{n,m}| \colon 0 \le m < 2^n\}$) tends to 0 as $n \to \infty$, the Haar family forms an orthonormal basis in the space $L^2(J) \ominus \{\text{const}\}$.

Definition. Fix a real p, $1<p<\infty$, and let $p'=\frac{p}{p-1}$, $p^*=\max\{p,p'\}$. Introduce the following domain in $\mathbb{R}^3$:

$$\Omega=\Omega(p)=\{x=(x_1,x_2,x_3):\ x_3\ge 0,\ |x_1|^p\le x_3\}\,.$$

For a fixed partition $\mathcal{I}$ of an interval J we define two function on this domain

$$\mathbf{B}_{\max}(x)=\mathbf{B}_{\max}(x;p)=\sup_{f,\,g}\big\{\langle|g|^p\rangle_{_J}\big\},$$
$$\mathbf{B}_{\min}(x)=\mathbf{B}_{\min}(x;p)=\inf_{f,\,g}\big\{\langle|g|^p\rangle_{_J}\big\}.$$

where the supremum is taken over all functions f, g from $L^p(J)$ such that $\langle f\rangle_{_J}=x_1$, $\langle g\rangle_{_J}=x_2$, $\langle|f|^p\rangle_{_J}=x_3$, and $|(f,h_{_I})|=|(g,h_{_I})|$. We shall refer to any such pair of functions f, g as to an admissible pair. When $|(f,h_{_I})|=|(g,h_{_I})|$ happens for all dyadic intervals inside J **we call g a Martingale transform of** f. We shall call $\mathbf{B}_{\max}(x)$ (and $\mathbf{B}_{\max}(x)$) the Bellman functions of the problem of finding the best constant for the Martingale transform inequality:

$$|\langle g\rangle_{_J}|\le|\langle f\rangle_{_J}|\Rightarrow\langle|g|^p\rangle_J\le C(p)\langle|f|^p\rangle_J\,. \tag{13.3}$$

This best constant was found by Burkholder: $C(p)=(p^*-1)^p$, $p^*:=\max(p,\frac{p}{p-1})$.
Remark 1. It is amazing that there is no proof that would find this $C(p)$ without finding the function of 3 variables $\mathbf{B}_{\max}(x)$ or some of its relatives (like, for example, u_p from (13.1)).

Remark 2. Burkholder proved that the functions $\mathbf{B}$ do not depend on the initial interval J and on a specific choice of its partition. Below we work only with dyadic partitions.

Remark 3. In the case $p=2$ the Bellman function are evident:

$$\mathbf{B}_{\max}(x)=\mathbf{B}_{\min}(x)=x_2^2+x_3-x_1^2\,.$$

Indeed, since

$$\|f\|_2^2=|J|x_3=|J|x_1^2+\sum_{I\in\mathcal{I}}|(f,h_{_I})|^2\,,$$

we have

$$\begin{aligned}\langle|g|^2\rangle_{_J}&=\frac{1}{|J|}\|g\|_2^2=x_2^2+\frac{1}{|J|}\sum_{I\in\mathcal{I}}|(g,h_{_I})|^2\\&=x_2^2+\frac{1}{|J|}\sum_{I\in\mathcal{I}}|(f,h_{_I})|^2=x_2^2+x_3-x_1^2\,.\end{aligned}$$

Define the following function on $\mathbb{R}^2_+ = \{z = (z_1, z_2)\colon z_i > 0\}$:

$$F_p(z_1,z_2)=\begin{cases}\big[z_1^p-(p^*-1)^p z_2^p\big], \text{ if } z_1\le (p^*-1)z_2\,,\\ p(1-\frac{1}{p^*})^{p-1}(z_1+z_2)^{p-1}\big[z_1-(p^*-1)z_2\big], \text{ if } z_1\ge (p^*-1)z_2\,.\end{cases} \tag{13.4}$$

Note for for $p = 2$ the expressions above are reduced to $F_2(z_1, z_2) = 2(z_1^2 - z_2^2)$.

14. The main result

Now we are ready to state the main result:

Theorem 32. *The equation $F_p(|x_1|, |x_2|) = F_p(x_3^{\frac{1}{p}}, \mathbf{B}^{\frac{1}{p}})$ determines implicitly the function $\mathbf{B} = \mathbf{B}_{\min}(x; p)$ and the equation $F_p(|x_2|, |x_1|) = F_p(\mathbf{B}^{\frac{1}{p}}, x_3^{\frac{1}{p}})$ determines implicitly the function $\mathbf{B} = \mathbf{B}_{\max}(x; p)$.*

Remark. The reader can take a look at formulae (5.23)–(5.27) on page 660 of [Bu1] and recognize that this is how Burkholder describes $\mathbf{B}_{\max}$. The same is true for $\mathbf{B}_{\min}$.

15. How to find Bellman functions

We start from deducing the main inequality for Bellman functions. Introduce new variables $y_1 = \frac{1}{2}(x_2 + x_1)$, $y_2 = \frac{1}{2}(x_2 - x_1)$, and $y_3 = x_3$. In terms of the new variables we define a function $\mathcal{M}$,

$$\mathcal{M}(y_1, y_2, y_3) = \mathbf{B}(x_1, x_2, x_3) = \mathbf{B}(y_1 - y_2, y_1 + y_2, y_3)\,,$$

on the domain

$$\Xi = \{y = (y_1, y_2, y_3)\colon\ y_3 \ge 0,\ |y_1 - y_2|^p \le y_3\}\,.$$

Since the point of the boundary $x_3 = |x_1|^p$ ($y_3 = |y_1 - y_2|^p$) occurs for the only constant test function $f = x_1$ (and therefore then $g = x_2$ is a constant function as well) we

$$\mathbf{B}(x_1, x_2, |x_1|^p) = |x_2|^p\,,$$

or

$$\mathcal{M}(y_1, y_2, |y_1 - y_2|^p) = |y_1 + y_2|^p\,. \tag{15.1}$$

Note that the function $\mathbf{B}$ is even with respect of x_1 and x_2, i.e.,

$$\mathbf{B}(x_1, x_2, x_3) = \mathbf{B}(-x_1, x_2, x_3) = \mathbf{B}(x_1, -x_2, x_3)\,.$$

It follows from the definition of $\mathbf{B}$ if we consider the test functions $\tilde f = -f$ for the first equality and $\tilde g = -g$ for the second one. For the function $\mathcal{M}$ this means that we have the symmetry with respect to the lines $y_1 = \pm y_2$

$$\mathcal{M}(y_1, y_2, y_3) = \mathcal{M}(y_2, y_1, y_3) = \mathcal{M}(-y_1, -y_2, y_3)\,. \tag{15.2}$$

Therefore, it is sufficient to find the function $\mathbf{B}$ in the domain

$$\Omega_+ = \Omega_+(p) = \{x = (x_1, x_2, x_3) :\ x_i \ge 0,\ |x_1|^p \le x_3\}\,, \tag{15.3}$$

or the function $\mathcal{M}$ in the domain

$$\Xi_+ = \{y = (y_1, y_2, y_3) \colon\ -y_1 \le y_2 \le y_1,\ y_3 \ge 0,\ (y_1 - y_2)^p \le y_3\}\,. \tag{15.4}$$

Then we get the solution in the whole domain by putting

$$\mathbf{B}(x_1, x_2, x_3) = \mathbf{B}(|\,x_1|, |\,x_2|, x_3)\,.$$

Due to the symmetry (15.2) we have the following boundary conditions on the "new part" of the boundary $\partial\Xi_+$:

$$\begin{aligned} \frac{\partial\mathcal{M}}{\partial y_1} &= \frac{\partial\mathcal{M}}{\partial y_2} \qquad \text{on the hyperplane } y_2 = y_1\,, \\ \frac{\partial\mathcal{M}}{\partial y_1} &= -\frac{\partial\mathcal{M}}{\partial y_2} \qquad \text{on the hyperplane } y_2 = -y_1\,. \end{aligned} \tag{15.5}$$

If we consider the family of test functions $\tilde f = \tau f$, $\tilde g = \tau g$ together with f and g we come to the following homogeneity condition

$$\mathbf{B}(\tau x_1, \tau x_2, \tau^p x_3) = \tau^p \mathbf{B}(x_1, x_2, x_3)\,,$$

or

$$\mathcal{M}(\tau y_1, \tau y_2, \tau^p y_3) = \tau^p \mathcal{M}(y_1, y_2, y_3)\,.$$

We shall use this property in the following form: take derivative with respect to τ and put $\tau = 1$

$$y_1 \frac{\partial\mathcal{M}}{\partial y_1} + y_2 \frac{\partial\mathcal{M}}{\partial y_2} + p y_3 \frac{\partial\mathcal{M}}{\partial y_3} = p\mathcal{M}(y_1, y_2, y_3)\,. \tag{15.6}$$

Let us fix two points $x^\pm \in \Omega$ such that $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$, for the corresponding points $y^\pm \in \Xi$ this means that either $y_1^+ = y_1^-$, or $y_2^+ = y_2^-$. Then for an arbitrarily small number $\varepsilon > 0$ by the definition of the Bellman function $\mathbf{B} = \mathbf{B}_{\max}$ there exist two couples of test functions $f^\pm$ and $g^\pm$ on the intervals $I^\pm$ such that $\langle f^\pm \rangle_{I^\pm} = x_1^\pm$, $\langle g^\pm \rangle_{I^\pm} = x_2^\pm$, $\langle |f^\pm|^p \rangle_{I^\pm} = x_3^\pm$, and $\langle |g^\pm|^p \rangle_{I^\pm} \ge \mathbf{B}(x^\pm) - \varepsilon$. On the interval $I = I^+ \cup I^-$ we define a pair of test functions f and g as follows

$f|I^\pm = f^\pm$, $g|I^\pm = g^\pm$. This is a pair of test functions that corresponds to the point $x = \alpha^+ x^+ + \alpha^- x^-$, where $\alpha^\pm = |I^\pm|/|I|$, because the property $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$ means $|(f, h_I)| = |(g, h_I)|$. This yields

$$\mathbf{B}(x) \geq \langle |g|^p \rangle_I = \alpha^+ \langle |g^+|^p \rangle_I^+ + \alpha^- \langle |g^-|^p \rangle_I^- \geq \alpha^+ \mathbf{B}(x^+) + \alpha^- \mathbf{B}(x^-) - \varepsilon \,.$$

Since ε is arbitrary we conclude

$$\mathbf{B}(x) \geq \alpha^+ \mathbf{B}(x^+) + \alpha^- \mathbf{B}(x^-) \,. \tag{15.7}$$

For the function $\mathbf{B} = \mathbf{B}_{\min}$ we can get in a similar way

$$\mathbf{B}(x) \leq \alpha^+ \mathbf{B}(x^+) + \alpha^- \mathbf{B}(x^-) \,. \tag{15.8}$$

Recall that this is not quite concavity (convexity) condition, because we have the restriction $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$. But in terms of the function $\mathcal{M}$

$$\mathcal{M}_{\max}(y) \geq \alpha^+ \mathcal{M}_{\max}(y^+) + \alpha^- \mathcal{M}_{\max}(y^-) \,,$$
$$\mathcal{M}_{\min}(y) \leq \alpha^+ \mathcal{M}_{\min}(y^+) + \alpha^- \mathcal{M}_{\min}(y^-) \,,$$

when either $y_1 = y_1^+ = y_1^-$, or $y_2 = y_2^+ = y_2^-$, we indeed have the concavity (convexity) of the function $\mathcal{M}$ with respect to y_2, y_3 under a fixed y_1, and with respect to y_1, y_3 under a fixed y_2.

Since the domain is convex, under the assumption that the function $\mathbf{B}$ are sufficiently smooth these conditions of concavity (convexity) are equivalent to the differential inequalities

$$\begin{pmatrix} \mathcal{M}_{y_1y_1} & \mathcal{M}_{y_1y_3} \\ \mathcal{M}_{y_3y_1} & \mathcal{M}_{y_3y_3} \end{pmatrix} \leq 0 \,, \quad \begin{pmatrix} \mathcal{M}_{y_2y_2} & \mathcal{M}_{y_2y_3} \\ \mathcal{M}_{y_3y_2} & \mathcal{M}_{y_3y_3} \end{pmatrix} \leq 0 \,, \quad \forall y \in \Xi \,, \tag{15.9}$$

for $\mathcal{M} = \mathcal{M}_{\max}$ (here $\mathcal{M}_{y_iy_j}$ stand for the partial derivatives $\dfrac{\partial^2 \mathcal{M}}{\partial y_i \partial y_j}$) and

$$\begin{pmatrix} \mathcal{M}_{y_1y_1} & \mathcal{M}_{y_1y_3} \\ \mathcal{M}_{y_3y_1} & \mathcal{M}_{y_3y_3} \end{pmatrix} \geq 0 \,, \quad \begin{pmatrix} \mathcal{M}_{y_2y_2} & \mathcal{M}_{y_2y_3} \\ \mathcal{M}_{y_3y_2} & \mathcal{M}_{y_3y_3} \end{pmatrix} \geq 0 \,, \quad \forall y \in \Xi \,, \tag{15.10}$$

for $\mathcal{M} = \mathcal{M}_{\min}$.

Extremal properties of the Bellman function requires for one of matrices in (15.9) and (15.10) to be degenerated. So we arrive at the Monge–Ampère equation:

$$\mathcal{M}_{y_iy_i} \mathcal{M}_{y_3y_3} = (\mathcal{M}_{y_iy_3})^2 \tag{15.11}$$

either for $i = 1$ or for $i = 2$. To find a candidate M for the role of the true Bellman function $\mathcal{M}$ we shall solve this equation. After finding this solution we shall prove that $M = \mathcal{M}$.

The method of solving homogeneous Monge–Ampère equation is described, for example, in [VaVo]. In particular we know that the solution of the Monge–Ampère equation has to be of the form

$$M = t_i y_i + t_3 y_3 + t_0\,, \tag{15.12}$$

it is linear along the lines (let us call them *extremal trajectories*)

$$y_i dt_i + y_3 dt_3 + dt_0 = 0\,. \tag{15.13}$$

One of the ends of the extremal trajectory has to be a point on the boundary $y_3 = |y_1 - y_2|^p$, where constant functions are the only test functions corresponding to these points. Denote this point by $U = (y_1, u, (y_1 - u)^p)$. Note that we write $(y_1 - u)^p$ instead of $|y_1 - u|^p$ because the domain Ξ_+ is under consideration. For the second end of the extremal trajectory we have four possibilities

1) it belongs to the same boundary $y_3 = (y_1 - y_2)^p$;
2) it is at infinity $(y_1, y_2, +\infty)$, i.e., the extremal lines goes parallel to the y_3-axis;
3) it belongs to the boundary $y_2 = y_1$;
4) it belongs to the boundary $y_2 = -y_1$.

The first possibility gives us no solution. Namely, we have the following

Theorem 33. *Function* $\mathbf{B}_{max}$ *cannot be equal to* $B(x) = M(y)$, *where* M *is the solution of the Monge–Ampère equation* (15.11) *such that one of its extremal trajectory is of type 1) above. The same claim holds for* $\mathbf{B}_{min}$.

To see that, the reader can think that we are in the plain $y_1 = \text{const}$ (the case $y_2 = const$ is totally the same). Suppose $\mathcal{M}_{\max}$ (or $\mathcal{M}_{\min}$) is linear on a segment, which connects two points of the "parabola" $y_3 = (y_1 - u)^p, -y_1 \le u \le y_1$ (see (15.4)). We will show now that the true Bellman functions (either $\mathcal{M}_{\max}$ or $\mathcal{M}_{\min}$) cannot be linear on a chord connecting two points of this part of the parabola $y_3 = (y_1 - y_2)^p$ (and simultaneously be equal to M, a solution of the Monge–Ampère equation (15.11)).

To check this it is sufficient verify that the test function $f(t) = \alpha + \beta H_I(t)$ cannot be an extremal function of our problem with the only exception $p = 2$, when the situation is trivial: $\mathbf{B}_{\max}(x) = \mathbf{B}_{\min}(x) = x_3 + x_2^2 - x_1^2$, and any test function is extremal. In fact, if $\mathbf{B}_{\max}$ (or $\mathbf{B}_{\min}$) is linear on a line which connects two points x^- and x^+, and (as we assumed for the sake of definitivity) $y_1^- = y_1^+ = y_1$, we consider the middle point x of the interval $[x^-, x^+]$. Let $x := (x_1, x_2, x_3)$, $a := \frac{1}{2}(x_1^+ - x_1^-)$. But we consider the case $x_1^+ + x_2^+ = y_1^+ = y_1^- = x_1^- + x_2^-$. Therefore, $a := \frac{1}{2}(x_2^- - x_2^+)$. Then

$$x_1^- = x_1 - a,\ x_1^+ = x_1 + a,\ x_2^- = x_2 + a,\ x_2^+ = x_2 - a\,.$$

Let $I = [0,1]$. We consider the first Haar function H_I (by definition it is equal to 1 on $[0, 1/2]$ and to -1 on $(-1/2, 1])$ and we consider

$$f^- = x_1^- = x_1 - a \text{ on } [0,1],\ g^- = x_2^- = x_2 + a \text{ on } [0,1],$$

$$f^+ = x_1^+ = x_1 + a \text{ on } [0,1],\ g^+ = x_2^+ = x_2 - a \text{ on } [0,1],$$

$$f = x_1 - aH_I,\ \ g = x_2 + aH_I\,.$$

Notice that g is the Martingale transform of f as defined above.

Let $x_3 := \langle |f|^p \rangle$. As before

$$y_1 = \frac{x_1 + x_2}{2}\,,\ y_2 = \frac{x_2 - x_1}{2}\,.$$

or

$$x_1 = y_1 - y_2\,,\ x_2 = y_1 + y_2\,.$$

Let $x_3^- := |x_1^-|^p\,,\ x_3^+ := |x_1^+|^p$

$$x := (x_1, x_2, x_3)\,,\ x^- := (x_1^-, x_2^-, x_3^-)\,,\ x^+ := (x_1^+, x_2^+, x_3^+)\,.$$

As we can consider only the case

$$x^-, x^+, x = \frac{x^- + x^+}{2} \in \text{ the First quarter} \times [0, \infty)\,, \tag{15.14}$$

(and we may think without the loss of generality that $a > 0$) we, thus, consider only the case when

$$a \le x_1, a \le x_2, a, x_1, x_2 \ge 0\,. \tag{15.15}$$

Also

$$y_2^- = y_2 + a\,,\ y_2^+ = y_2 - a\,,\ y_2 \text{ are of any sign}\,. \tag{15.16}$$

Obviously

$$\mathbf{B}_{\max}(x^-) = \mathbf{B}_{\min}(x^-) = x_3^- := |x_1^-|^p,\ \mathbf{B}_{\max}(x^+) = \mathbf{B}_{\min}(x^+) = x_3^+ := |x_1^+|^p\,.$$

And the extremal pairs of functions are (f^-, g^-), (f^+, g^+) correspondingly. The linearity of $\mathbf{B}_{\max}$ (or $\mathbf{B}_{\min}$) on the segment $[x^-, x^+]$ shows that the pair (f, g) just defined is an extremal pair for $x = \frac{1}{2}(x^- + x^+)$, and

$$\mathbf{B}_{\max}(x) = \langle |x_2 + aH_I|^p \rangle\,. \tag{15.17}$$

(Or the same (15.17) but with $\mathbf{B}_{\min}(x)$.) We want to bring (15.17) to contradiction. For this we will need to consider several cases.

Remark. The interesting feature of the reasoning below is that *it is not a reasoning by small perturbation of* (f, g) pair. We will perturb f and g, but by a *big* perturbation.

Consider ϕ_h equal to 1 on $[0, 1/2 - h] \cup [1 - h, 1]$ and to -1 on $(1/2 - h, 1 - h)$. We have

$$(\phi_h, H_I) = 1 - 4h\,.$$

We are interested in $\phi := \phi_{1/8}$. Then

$$(\phi, H_I) = 1/2\,,$$

and $\Phi = \phi - H_I$ has

$$(\Phi, H_I) = -1/2\,, (\Phi, H_J) = (\phi, H_J)$$

for all other dyadic J. So Φ is a martingale transform of ϕ. It is equal to 0 on $[0, 3/8] \cup [1/2, 7/8]$ and to ± 2 on two intervals $(7/8, 1)$, $(3/8, 1/2)$.

The case of $\mathbf{B}_{\max}$.

Put

$$\tilde f := x_1 - a\phi\,,\ \tilde g := x_2 + a\Phi\,.$$

Lemma 34. $\|f\|_p = \|\tilde f\|_p\,.$

Proof. Function ϕ is just 1/8 shift of H_I. So f and $\tilde f$ have the same distribution function. □

Lemma 35. *If* $p \ge 3$ *or* $1 < p < 2$*, then* $\|g\|_p < \|\tilde g\|_p\,.$

Proof. First we will prove an auxiliary lemma:

Lemma 36. *If $p \geq 3$ or $1 < p < 2$, and if $\alpha \in (0,1)$, then*

$$g(\alpha) := \frac{1}{4}((1+2\alpha)^p + |2\alpha - 1|^p) + \frac{3}{2} - ((1+\alpha)^p + (1-\alpha)^p) > 0\,.$$

Proof. We will do it for $p \geq 3$. For any $p > 1$ we have $g(0) = 0, g'(0) = 0$. Also $b = p - 2 \geq 1$, so

$$\frac{1}{p(p-1)} g''(\alpha) = ((1+2\alpha)^b - (1+\alpha)^b) - ((1-\alpha)^b - |1-2\alpha|^b)\,.$$

Consider first $\alpha \in (0, 1/2]$. Notice that $x \to x^b$ is convex, so the first bracket is definitely bigger than the second one. This is because the argument increases by α in both brackets if $\alpha \in (0, 1/2]$. If $\alpha \in (1/2, 1)$ the increase of argument in the first bracket is still α, but it is $2 - 3\alpha$ in the second one. This is always a smaller increase, and we again get that the first bracket is bigger than the second one for the convex function. So we proved lemma for $p \geq 3$. The same proof works for $1 < p < 2$.

□

Now denote $\alpha := \frac{a}{x_2}$ and write

$$2\|\tilde{g}\|_p^p = \frac{1}{4}((x_2 + 2a)^p + |x_2 - 2a|^p) + \frac{3}{2}x_2^p = x_2^p(\frac{1}{4}((1+2\alpha)^p + |1-2\alpha|^p) + \frac{3}{2})\,.$$

$$2\|g\|_p^p = x_2^p((1+\alpha)^p + (1-\alpha)^p)\,.$$

Therefore Lemma 36 gives the desired inequality $\|\tilde{g}\|_p^p - \|g\|_p^p > 0$.

So our Lemma 35 is also proved.

□

Remark. Notice that

$$\text{for } 2 < p < 3 \text{ we have: } g''(\alpha) < 0\,, \alpha \in (0, 1/2)\,. \tag{15.18}$$

Now we are left to work with $2 < p < 3$ case. Using exactly the same reasoning as in Lemma 35–only with (15.18)–we get the following

Lemma 37. *Let $2 < p < 3$ and $0 < a \leq \frac{x_1}{2}$, then $\|x_1 - a\Phi\|_p < \|x_1 - aH_I\|_p$.*

So in the case $0 < a \leq \frac{x_1}{2}$ we can finish our reasoning as follows. Obviously, for any x,

$$t \to \|x - t\Phi\|_p^p = \frac{1}{8}(|x - 2t|^p + |x + 2t|^p) + \frac{3}{4}|x|^p\,, \text{ is increasing in } t > 0\,, \tag{15.19}$$

and

$$t \to \|x - tH_I\|_p^p = \frac{1}{2}(|x-t|^p + |x+t|^p) \text{ is increasing in } t > 0\,. \tag{15.20}$$

And the same with "minus" replaced by "plus". This is just because $x \to |x|^p$ is convex. Using Lemma 37 we can now choose $A > a$ such that

$$\|x_1 - A\Phi\|_p = \|x_1 - aH_I\|_p\,, \|x_2 + A\Phi\|_p > \|x_2 + aH_I\|_p\,. \tag{15.21}$$

Now we have almost as before (but with interchanged ϕ and Φ)

$$f = x_1 - aH_I\,, g = x_2 + aH_I\,, \tilde{f} := x_1 - A\Phi\,, \tilde{g} := x_2 + A\phi\,.$$

Then (15.21) means that $\|f\|_p = \|\tilde{f}\|_p$ but $\|\tilde{g}\|_p = \|x_2 + AH_I\|_p > \|g\|_p$, and this is exactly what we need to disprove that (f,g) is an extremal pair for $B_{\max}$.

For $\mathbf{B}_{\max}$ we almost finished the proof of Theorem 33. We are left with the case $\frac{x_1}{2} < a \le x_1$ (that is $\alpha(x_1,a) = a/x_1 \in (1/2,1]$), $2 < p < 3$. In this range of (a, x_1) Lemma 37 may be not true anymore. Namely, $\|x_1 - a\Phi\|_p < \|x_1 - aH_I\|_p$ might not hold anymore. And sometimes it does not hold, for example, if $p \approx 3, p < 3$, $\alpha = 1$ we get $\frac{1}{4}(3^p+1) + \frac{3}{2} > 2^p$ because for $p = 3$ this becomes $34 > 32$.

So let $\frac{x_1}{2} < a \le x_1$ (that is $\alpha(x_1,a) = a/x_1 \in (1/2,1]$), $2 < p < 3$, and suppose that we have a chord connecting two points of our parabola, and this chord represents a characteristic (we called them also *extremal trajectories*) of the corresponding Monge-Ampère equation (here we use our assumption that $B_{\max}(x) = B(x) = M(y)$, where M solves (15.11)).

Let us consider the crescent "below" the chord $L := [x^-, x^+]$ and "above" the parabola. It should be be filled in by chords on which $M(y) = B(x)$ are linear (this is the property of the solutions of the homogeneous Monge–Ampère equation expressed in Pogorelov's theorem, see [Pog]), which can have with our chord L at most one common point (one of the ends). Two cases may happen: 1) there is a chord in our crescent which has no common point with L, 2) chords form the fan exiting from one end point (say) x^- of L. In the first case we necessarily will have a chord $\ell := [z^-, z^+]$ such that

$$0 < c := \frac{|z_1^+ - z_1^-|}{2} \le \frac{z}{2}\,,\; z := \frac{z_1^+ + z_1^-}{2}\,. \tag{15.22}$$

But this is the case we draw to a contradiction just a second ago.

Suppose now the fan occurs. Again if 2a) $x_1^- = x_1 - a > 0$, the same ℓ can be found. Finally, if 2b) $x_1^- = x_1 - a = 0$, then we can of course have a very small chord $\ell := [z^-, z^+]$, $z^- = x^-, z^+ \to x^-$, and again we can achieve (15.22). In this case the second claim of (15.22) is achieved with equality (clearly if $x_1^- = 0$ then $x_1^+ \neq 0$, otherwise x^- and x^+ cannot both lie on the parabola without being the same point).

Theorem 33 is completely proved for the case $\mathbf{B}_{\max}$.

The case of $\mathbf{B}_{\min}$.

We will use the same Φ and ϕ as in the case of $\mathbf{B}_{\max}$ above, but we will build $\tilde{f}, \tilde{g}$ differently. More precisely, our choice will be dual to the one above. Let first $2 < p < 3$. Then we use Lemma 37 to conclude that if $0 < a \leq \frac{x_2}{2}$ then

$$\|x_2 + a\Phi\|_p < \|x_2 + a\phi\|_p \,. \tag{15.23}$$

Then we put

$$\tilde{f} := x_1 - a\phi,\ \tilde{g} := x_2 + a\Phi \,.$$

We obtain immediately

$$\|\tilde{f}\|_p = \|f\|_p$$

just by looking at the fact that their distribution functions are the same, and we obtain

$$\|\tilde{g}\|_p < \|g\|_p$$

by (15.23). So $\mathbf{B}_{\min}$ cannot be linear on such a chord.

If we have $\frac{x_2}{2} < a \leq x_2$, we reason as around (15.22). Namely, in this case, if $\mathbf{B}_{\min}$ is equal to $B(x) = M(y)$, where M is a solution of a Monge–Ampère equation (15.11), then there is another small chord $[X^-, X^+]$ such that

$$0 < c := \frac{|X_2^+ - X_2^-|}{2} \leq \frac{X}{2} \,,\, X := \frac{X_2^+ + X_2^-}{2} \,, \tag{15.24}$$

and we can repeat the previous considerations with the new data.

Now we are in the situation when either $p \geq 3$ or $1 < p < 2$. By Lemma 35 we have then

$$\|x_1 - a\Phi\|_p > \|x_1 - a\phi\|_p = \|x_1 - aH_I\|_p = \|f\|_p > |x_1|^p \,. \tag{15.25}$$

By (15.19) we can choose $0 < b < a$ such that

$$\|x_1 - b\Phi\|_p = \|x_1 - aH_I\|_p = \|f\|_p \,.$$

Using (15.20) we obtain that

$$\|x_2 + b\phi\|_p = \|x_2 + bH_I\|_p < \|x_2 + aH_I\|_p = \|g\|_p \,.$$

Now put

$$\tilde f := x_1 - b\Phi,\ \tilde g := x_2 + b\phi \,.$$

We have just seen that

$$\|\tilde f\|_p = \|f\|_p,\ \|\tilde g\|_p < \|g\|_p \,.$$

So in this range of p $\mathbf{B}_{\min}$ cannot be linear on a chord connecting two points of the parabola. Theorem 33 is completely proved.

Now we check the second possibility among the possibilities 1)–4) listed right before Theorem 33. Since the extremal line is parallel to the y_3-axis, the Bellman function has to be of the form

$$M(y) = A(y_1, y_2) + C(y_1, y_2)y_3 \,.$$

Any pair of inequalities both (15.9) and (15.10) implies $M_{y_iy_i}M_{y_3y_3} - (M_{y_iy_3})^2 \geq 0$. Since $M_{y_3y_3} = 0$, this yields $M_{y_iy_3} = \frac{\partial C}{\partial y_i} = 0$, i.e., C is a constant. From the boundary condition (15.1) we get

$$A(y_1, y_2) + C(y_1 - y_2)^p = (y_1 + y_2)^p \,,$$

whence

$$A(y_1, y_2) = (y_1 + y_2)^p - C(y_1 - y_2)^p \,,$$

and

$$M(y) = (y_1 + y_2)^p + C(y_3 - (y_1 - y_2)^p) \,, \tag{15.26}$$

or

$$B(x) = x_2^p + C(x_3 - x_1^p) \,. \tag{15.27}$$

Let us note that this solution cannot satisfy necessary conditions in the whole domain Ξ_+ except the case $p = 2$. The constant C must be positive (otherwise the extremal lines cannot tend to infinity along y_3-axes, because M must be a nonnegative function). Therefore, the straight line

$$y_1 + y_2 = C^{\frac{1}{p-2}}(y_1 - y_2),\ \text{or } x_2 = C^{\frac{1}{p-2}}y_1$$

splits Ξ_+ in two subdomains, in one of which the derivatives

$$\frac{\partial^2 M}{\partial y_1^2} = \frac{\partial^2 M}{\partial y_2^2} = p(p-1)\Big((y_1+y_2)^{p-2} - C(y_1-y_2)^{p-2})\Big)$$

is positive (i.e., it could be a candidate for $\mathbf{B}_{\min}$), and in another one is negative (i.e., it could be a candidate for $\mathbf{B}_{\max}$).

Thus, this simple solution cannot give us the whole Bellman function and we need to continue the consideration of the possibilities 3) and 4) (listed right before Theorem 33). Till now we have not fixed which of two matrices in (15.9) or in (15.10) is degenerated, i.e., what is i in the Monge–Ampère equation (15.11), because for the vertical extremal lines both these equations are fulfilled. Now, when considering possibility 3) or 4), we need to investigate separately both Monge–Ampère equations (15.11). We shall refer to these cases as $3_{i)}$ and $4_{i)}$.

Let us start with simultaneous consideration of the cases $3_{1)}$ (we recall that this means that y_2 is fixed). and $4_{1)}$. We look for a function

$$M = t_1y_1 + t_3y_3 + t_0$$

on the domain Ξ_+, which is linear along the extremal lines

$$y_1dt_1 + y_3dt_3 + dt_0 = 0\,.$$

Now one end point of our extremal line $V = (v, y_2, (v-y_2)^p)$ belongs to the boundary $y_2 = |y_1 - y_2|^p$ and the second end point $W = (|y_2|, y_2, w)$ is on the boundary $y_1 = |y_2|$, where we have boundary condition (15.5). Due to the symmetry (15.2), on the boundary $y_1 = y_2$ (this means that our fixed $y_2 \geq 0$) we have

$$\frac{\partial M}{\partial y_2} = \frac{\partial M}{\partial y_1} = t_1\,,$$

and

$$\frac{\partial M}{\partial y_2} = -\frac{\partial M}{\partial y_1} = -t_1\,,$$

on the boundary $y_1 = -y_2$ (this means that our fixed $y_2 < 0$). In both cases

$$y_2\frac{\partial M}{\partial y_2} = y_1\frac{\partial M}{\partial y_1} = |y_2|t_1\,,$$

and therefore (15.6) and (15.12) imply

$$2t_1|y_2| + pwt_3 = pM(W) = pt_1|y_2| + pwt_3 + pt_0\,,$$

whence

$$t_0 = (\frac{2}{p} - 1)t_1|y_2|\,.$$

This gives the formula for $t_0(t_1)$ (remember that y_2 is fixed as we consider the cases 3_1), 4_1) now). Thus, we get

$$M(y) = \Big[y_1 + (\frac{2}{p} - 1)|y_2|\Big]t_1 + y_3t_3\,. \tag{15.28}$$

Since $dt_0 = \big(\frac{2}{p}-1\big)|y_2|\,dt_1$, the equation of the extremal trajectories (15.13) takes the form

$$\Big[y_1 + (\frac{2}{p} - 1)|y_2|\Big]dt_1 + y_3\,dt_3 = 0\,, \tag{15.29}$$

and we can rewrite (15.28) as follows

$$M(y) = \Big(t_3 - t_1\frac{dt_3}{dt_1}\Big)y_3\,.$$

We see that the expression $M(y)/y_3$ is constant along the trajectory and we can find it evaluating at the point V, where the boundary condition (15.1) is known:

$$M(y) = \left(\frac{v + y_2}{v - y_2}\right)^p y_3\,, \tag{15.30}$$

where $v = v(y_1, y_2, y_3)$ satisfies the following equation:

$$\frac{y_1 + \big(\frac{2}{p} - 1\big)|y_2|}{y_3} = \frac{v + \big(\frac{2}{p} - 1\big)|y_2|}{(v - y_2)^p}\,, \tag{15.31}$$

because the point $V = (v, y_2, (v - y_2)^p)$ is on the extremal line (15.29). We even shall not check under what conditions equation (15.31) has a solution and when it is unique. Later we show that in any case the function M we have found cannot be the Bellman function we are interested in, because neither condition (15.9) nor (15.10) can be fulfilled: the matrix $\{M_{y_iy_j}\}_{i,j=2,3}$ is neither negative definite nor positive definite. We postpone this verification, because the calculations of the sign of the Hessian matrices is the same for this solution and another solution of the Monge–Ampère equation that supplies us with the true Bellman function. And these calculations will be made simultaneously a bit later. And now we only rewrite our solution in an implicit form more convenient for calculation.

We introduce

$$\omega := \left(\frac{M(y)}{y_3}\right)^{\frac{1}{p}}\,, \tag{15.32}$$

then (15.30) yields

$$v = \frac{\omega + 1}{\omega - 1}\,y_2\,. \tag{15.33}$$

Since $v \geq 0$ (in fact, recall that we consider now only y: $y_1 \geq |y_2|$ domain now, and that v is just the first coordinate of the point $V = (v, y_2, (v-y_2)^p$ in this domain), we have

$$\operatorname{sign} y_2 = \operatorname{sign}(\omega - 1)\,. \tag{15.34}$$

After substitution of (15.33) in (15.31) we get

$$\Big(\frac{2y_2}{\omega-1}\Big)^p\Big[y_1 + \big(\frac{2}{p}-1\big)|y_2|\Big] = y_3\Big[\frac{\omega+1}{\omega-1}y_2 + \big(\frac{2}{p}-1\big)|y_2|\Big]$$

or

$$2^p|y_2|^{p-1}\big[py_1 + (2-p)|y_2|\big] = y_3|\omega-1|^{p-1}\big[(\omega+1)p + (2-p)|\omega-1|\big]$$

For the case 3_1) we have $y_2 > 0$ (i.e., $x_2 > x_1$, we look for $\omega > 1$ or $B > y_3$) and the latter equation can be rewritten in the initial coordinates as follows

$$(x_2 - x_1)^{p-1}\big[x_2 + (p-1)x_1\big] = (B^{\frac{1}{p}} - x_3^{\frac{1}{p}})^{p-1}\big[B^{\frac{1}{p}} + (p-1)x_3^{\frac{1}{p}}\big]\,.$$

For the case 4_1) we have $y_2 < 0$ (i.e., $x_2 < x_1$, we look for $\omega < 1$ or $B < y_3$) and the equation takes the form

$$(x_1 - x_2)^{p-1}\big[x_1 + (p-1)x_2\big] = (x_3^{\frac{1}{p}} - B^{\frac{1}{p}})^{p-1}\big[x_3^{\frac{1}{p}} + (p-1)B^{\frac{1}{p}}\big]$$

Introduce the following function

$$G(z_1, z_2) = (z_1 + z_2)^{p-1}\big[z_1 - (p-1)z_2\big]$$

defined on the half-plane $z_1 + z_2 \geq 0$. Then in the case 3_1) we have the relation

$$G(x_2, -x_1) = G(B^{\frac{1}{p}}, -x_3^{\frac{1}{p}})\,, \tag{15.35}$$

or

$$G(y_2 + y_1, y_2 - y_1) = y_3 G(\omega, -1)\,.$$

In the case 4_1) we have

$$G(x_1, -x_2) = G(x_3^{\frac{1}{p}}, -B^{\frac{1}{p}})\,, \tag{15.36}$$

or

$$G(y_1 - y_2, -y_1 - y_2) = y_3 G(1, -\omega)\,.$$

Now we have to consider the Monge-Ampère equation (15.11) in the cases 3_2) and 4_2). This means that we fix y_1 now. Let us begin with the cases 3_2), when an extremal line starts at a point $U=(y_1,u,(y_1-u)^p)$ on our parabola and ends at a point $W=(y_1,y_1,w)$. Again, the symmetry condition at the point W is

$$\frac{\partial M}{\partial y_1}=\frac{\partial M}{\partial y_2}=t_2\,,$$

and the homogeneity condition (15.6) plus condition (15.12) at W yield

$$2y_1t_2+pwt_3=pM(W)=py_1t_2+pwt_3+pt_0\,,$$

whence

$$t_0=\big(\frac2p-1\big)y_1t_2\,,$$

and therefore

$$M(y)=\Big[y_2+\big(\frac2p-1\big)y_1\Big]t_2+y_3t_3\,. \tag{15.37}$$

Since $dt_0=\big(\frac2p-1\big)y_1\,dt_2$, the equation of the extremal trajectories takes the form

$$\Big[y_2+\big(\frac2p-1\big)y_1\Big]dt_2+y_3\,dt_3=0\,, \tag{15.38}$$

and we can rewrite (15.37) as follows

$$M(y)=\Big(t_3-t_2\frac{dt_3}{dt_2}\Big)y_3\,.$$

Again, from here the expression $M(y)/y_3$ is constant along the trajectory and we can find it evaluating at the point U, where we the boundary condition (15.1) is known:

$$M(y)=\left(\frac{y_1+u}{y_1-u}\right)^p y_3\,, \tag{15.39}$$

where $u=u(y_1,y_2,y_3)$ can be found from (15.38):

$$\frac{y_2+\big(\frac2p-1\big)y_1}{y_3}=\frac{u+\big(\frac2p-1\big)y_1}{(y_1-u)^p}\,. \tag{15.40}$$

We see that if our extremal line starts at point $U=(y_1,u,(y_1-u)^p$ on our parabola $u=-\big(\frac2p-1\big)y_1$, then $y_2=-\big(\frac2p-1\big)y_1=u=$ const, i.e., it is a line parallel to the x_3-axes. This means that no extremal line that ends at the points of the boundary $y_1=y_2$ can intersect the plane $y_2=-\big(\frac2p-1\big)y_1$. This follows from the property that extremal trajectories do not intersect. Therefore, the starting points U with $u\le-\big(\frac2p-1\big)y_1$ cannot be acceptable for the case under consideration (since these trajectories do not intersect the plane $y_2=-\big(\frac2p-1\big)y_1$, they cannot have the second end point on $y_2=y_1$, see Fig. 1).

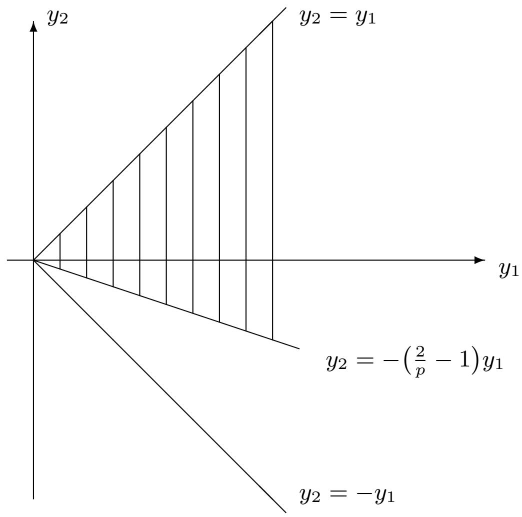

FIGURE 1. Acceptable sector for the case 3_2).

Let us check that equation (15.40) has exactly one solution $u = u(y_1, y_2, y_3)$ in the sector $-\big(\frac{2}{p} - 1\big)y_1 < y_2 < y_1$. Indeed, the function

$$u \mapsto y_3\Big[u + \big(\frac{2}{p} - 1\big)y_1\Big] - (y_1 - u)^p\Big[y_2 + \big(\frac{2}{p} - 1\big)y_1\Big]$$

is monotonously increasing for $u < y_1$ and it has the negative value $-\big(\frac{2}{p}y_1\big)^p\Big[y_2 + \big(\frac{2}{p} - 1\big)y_1\Big]$ at the point $u = -\big(\frac{2}{p} - 1\big)y_1$ and the positive value $\frac{2}{p}y_1y_3$ at the point $u = y_1$.

Now we rewrite the solution (15.39) in an implicit form using notations (15.32): $\omega := \Big(\frac{M(y)}{y_3}\Big)^{\frac{1}{p}}$. From (15.39) we have

$$u = \frac{\omega - 1}{\omega + 1}\, y_1\,, \tag{15.41}$$

therefore, from (15.40) we obtain

$$2^{-p}y_3(\omega + 1)^{p-1}[p(\omega - 1) + (2 - p)(\omega + 1)] = y_1^{p-1}\big[py_2 + (2 - p)y_1\big]$$

or

$$2^{-p+1}y_3(\omega + 1)^{p-1}(\omega - p + 1) = y_1^{p-1}\big[py_2 + (2 - p)y_1\big]\,,$$

which is (using again notations (15.32): $\omega := \Big(\frac{M(y)}{y_3}\Big)^{\frac{1}{p}}$)

$$(B^{\frac{1}{p}} + x_3^{\frac{1}{p}})^{p-1}\big[B^{\frac{1}{p}} - (p - 1)x_3^{\frac{1}{p}}\big] = (x_1 + x_2)^{p-1}\big[x_2 - (p - 1)x_1\big]\,.$$

In terms of function G this can be rewritten as follows

$$G(x_2,x_1)=G(B^{\frac1p},x_3^{\frac1p})\,,$$

or

$$G(y_1+y_2,y_1-y_2)=y_3G(\omega,1)\,.$$

It remains to examine the possibility 4_2). Assume that an extremal line starts at a point $U=(y_1,u,(y_1-u)^p)$ and ends at a point $W=(y_1,-y_1,w)$. Again, the homogeneity property (15.6) at the point W and the symmetry $\frac{\partial M}{\partial y_1}=-\frac{\partial M}{\partial y_2}=-t_2$ yield

$$-2y_1t_2+pwt_3=pM(W)=-py_1t_2+pwt_3+pt_0\,,$$

whence

$$t_0=\big(1-\frac2p\big)y_1t_2\,,$$

and therefore

$$M(y)=\Big[y_2+(1-\frac2p)\,y_1\Big]t_2+y_3t_3\,. \tag{15.42}$$

Since $dt_0=(1-\frac2p)y_1\,dt_2$, the equation of the extremal trajectories takes the form

$$\Big[y_2+(1-\frac2p)\,y_1\Big]dt_2+y_3\,dt_3=0\,, \tag{15.43}$$

and we can rewrite (15.42) as follows

$$M(y)=\Big(t_3-t_2\frac{dt_3}{dt_2}\Big)y_3\,.$$

Again, the expression $M(y)/y_3$ is constant along the trajectory and from the boundary condition (15.1) we get the same expression

$$M(y)=\left(\frac{y_1+u}{y_1-u}\right)^py_3\,. \tag{15.44}$$

Now $u=u(y_1,y_2,y_3)$ is a solution of the equation

$$\frac{y_2-\big(\frac2p-1\big)y_1}{y_3}=\frac{u-\big(\frac2p-1\big)y_1}{(y_1-u)^p} \tag{15.45}$$

that we get from (15.43). As before, we get trajectories ending at the plane $y_2=-y_1$ not in the whole domain Ξ_+, but only in the sector $-y_1<y_2<\big(\frac2p-1\big)y_1$ (see Fig. 2), and equation (15.45) has a unique solution for every point from this sector. As before, relation (15.41) allows us to rewrite the equation of extremal trajectories (15.45) as an implicite expression for ω (and hence for M):

$$G(x_1,x_2)=G(x_3^{\frac1p},B^{\frac1p})\,,$$

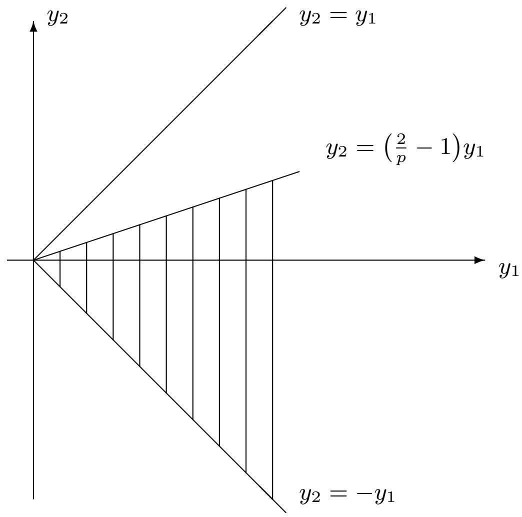

FIGURE 2. Acceptable sector for the case 4_2).

or

$$G(y_1 - y_2, y_1 + y_2) = y_3 G(1, \omega)\,.$$

Now we start the verification which of the obtained solutions satisfies conditions (15.9) or (15.10). We need to calculate $D_i := M_{y_i y_i} M_{y_3 y_3} - M^2_{y_i y_3}$, $i = 1, 2$, in four cases

3_1) $G(y_1 + y_2, -y_1 + y_2) = y_3 G(\omega, -1)$;
4_1) $G(y_1 - y_2, -y_1 - y_2) = y_3 G(1, -\omega)$;

$$3_2)\ G(y_1 + y_2, y_1 - y_2) = y_3 G(\omega, 1); \tag{15.46}$$

$$4_2)\ G(y_1 - y_2, y_1 + y_2) = y_3 G(1, \omega), \tag{15.47}$$

where $M = y_3 \omega^p$. In all situations we have a relation of the form

$$\Phi(\omega) = \frac{H(y_1, y_2)}{y_3}\,.$$

Till some moment in the future we will not specify the expression for Φ and H, as well as for their derivatives, and plug in the specific expression only in the final result after numerous cancellation. In particular, we introduce

$$R_1 = R_1(\omega) := \frac{1}{\Phi'} \quad \text{and} \quad R_2 = R_2(\omega) := R_1' = -\frac{\Phi''}{\Phi'^2}.$$

We would like to mention here that this idea, allowing us to make calculation shorter, is taken from the original paper of Burkholder [Bu1].

First of all we calculate the partial derivatives of ω:

$$\Phi'\omega_{y_3} = -\frac{H}{y_3^2} \qquad \Longrightarrow \qquad \omega_{y_3} = -\frac{R_1 H}{y_3^2}\,,$$
$$\Phi'\omega_{y_i} = \frac{H_{y_i}}{y_3} \qquad \Longrightarrow \qquad \omega_{y_i} = \frac{R_1 H_{y_i}}{y_3} = \frac{R_1 H'}{y_3}, \qquad i = 1, 2\,.$$

Here and further we shall use notation H' for any partial derivative H_{y_i}, $i = 1, 2$. This cannot cause misunderstanding because only one i participate in calculation of Hessian determinants D_i. Moreover, we shall not mention anymore that the index i can take two values either $i = 1$ or $i = 2$.

$$\omega_{y_3 y_3} = -\frac{R_2\omega_{y_3} H}{y_3^2} + 2\frac{R_1 H}{y_3^3} = \frac{R_1 H}{y_3^4}(R_2 H + 2y_3)\,,$$
$$\omega_{y_3 y_i} = -\frac{R_2\omega_{y_i} H}{y_3^2} - \frac{R_1 H'}{y_3^2} = -\frac{R_1 H'}{y_3^3}(R_2 H + y_3)\,,$$
$$\omega_{y_i y_i} = \frac{R_2\omega_{y_i} H'}{y_3} + \frac{R_1 H''}{y_3} = \frac{R_1}{y_3^2}(R_2(H')^2 + y_3 H'')\,.$$

Now we pass to the calculation of derivatives of $M = y_3\omega^p$:

$$M_{y_3} = p y_3 \omega^{p-1}\omega_{y_3} + \omega^p\,,$$
$$M_{y_i} = p y_3 \omega^{p-1}\omega_{y_i}\,;$$

$$\begin{aligned} M_{y_3 y_3} &= p y_3 \omega^{p-1}\omega_{y_3 y_3} + 2p\omega^{p-1}\omega_{y_3} + p(p-1) y_3 \omega^{p-2}\omega_{y_3}^2 \\ &= \frac{p\omega^{p-2} R_1 H^2}{y_3^3}[\omega R_2 + (p-1) R_1]\,, \qquad (15.48) \\ M_{y_3 y_i} &= p y_3 \omega^{p-1}\omega_{y_3 y_i} + p\omega^{p-1}\omega_{y_i} + p(p-1) y_3 \omega^{p-2}\omega_{y_3}\omega_{y_i} \\ &= -\frac{p\omega^{p-2} R_1 H H'}{y_3^2}[\omega R_2 + (p-1) R_1]\,, \\ M_{y_i y_i} &= p y_3 \omega^{p-1}\omega_{y_i y_i} + p(p-1) y_3 \omega^{p-2}\omega_{y_i}^2 \\ &= \frac{p\omega^{p-2} R_1}{y_3}\left([\omega R_2 + (p-1) R_1](H')^2 + \omega y_3 H''\right)\,. \end{aligned}$$

This yields

$$D_i = M_{y_3 y_3} M_{y_i y_i} - M_{y_3 y_i}^2 = \frac{p^2\omega^{2p-3} R_1^2 H^2 H''}{y_3^3}[\omega R_2 + (p-1) R_1]\,. \qquad (15.49)$$

Notice, that H' disappeared completely.

Now we need to calculate second derivatives of

$$H(y_1, y_2) = G(\alpha_1 y_1 + \alpha_2 y_2, \beta_1 y_1 + \beta_2 y_2),$$

where $\alpha_i, \beta_i = \pm 1$. And

$$\begin{aligned} H'' &= \frac{\partial^2}{\partial y_i^2} G(\alpha_1 y_1 + \alpha_2 y_2, \beta_1 y_1 + \beta_2 y_2) \\ &= \alpha_i^2 G_{z_1 z_1} + 2\alpha_i \beta_i G_{z_1 z_2} + \beta_i^2 G_{z_2 z_2} \\ &= G_{z_1 z_1} + G_{z_2 z_2} \pm 2G_{z_1 z_2} , \end{aligned}$$

where the "+" sign has to be taken if the coefficients in front of y_i are equal and the "−" sign in the opposite case.

The derivatives of G are simple:

$$\begin{aligned} G_{z_1} &= p(z_1 + z_2)^{p-2} \big[z_1 - (p-2) z_2 \big] , \\ G_{z_2} &= -p(p-1) z_2 (z_1 + z_2)^{p-2} ; \end{aligned}$$

$$\begin{aligned} G_{z_1 z_2} &= p(p-1)(z_1 + z_2)^{p-3} \big[z_1 - (p-3) z_2 \big] , \\ G_{z_1 z_2} &= -p(p-1)(p-2) z_2 (z_1 + z_2)^{p-3} , \\ G_{z_2 z_2} &= -p(p-1)(z_1 + z_2)^{p-3} \big[z_1 + (p-1) z_2 \big] . \end{aligned}$$

Note that $G_{z_1 z_1} + G_{z_2 z_2} = 2G_{z_1 z_2}$, and therefore, $H'' = 4G_{z_1 z_2}$ if $\alpha_i = \beta_i$ and $H'' = 0$ if $\alpha_i = -\beta_i$. The first case occurs for $H_{y_2 y_2}$ in cases 3_1), 4_1) and for $H_{y_1 y_1}$ in cases 3_2), 4_2). The second case occurs for $H_{y_1 y_1}$ in cases 3_1), 4_1) and for $H_{y_2 y_2}$ in cases 3_2), 4_2). In fact, we know that the equality $D_i = 0$ has to be fulfilled in the cases 3_i) and 4_i), because it is just the Monge–Ampère equation we have been solving.

So we have

3_1) $z_1 = y_1 + y_2$,

$z_2 = -y_1 + y_2$, $\quad G_{z_1 z_2} = p(p-1)(p-2)(y_1 - y_2)(2y_2)^{p-3}$,

4_1) $z_1 = y_1 - y_2$,

$z_2 = -y_1 - y_2$, $\quad G_{z_1 z_2} = p(p-1)(p-2)(y_1 + y_2)(-2y_2)^{p-3}$,

3_2) $z_1 = y_1 + y_2$,

$z_2 = y_1 - y_2$, $\quad G_{z_1 z_2} = -p(p-1)(p-2)(y_1 - y_2)(2y_1)^{p-3}$,

4_2) $z_1 = y_1 - y_2$,

$z_2 = y_1 + y_2$, $\quad G_{z_1 z_2} = -p(p-1)(p-2)(y_1 + y_2)(2y_1)^{p-3}$.

In the first pair of cases we have $\operatorname{sign} G_{z_1 z_2} = \operatorname{sign} H'' = \operatorname{sign}(p-2)$ and the opposite sign in the second pair of cases. In the first pair of cases we have that this sign is the sign of the Hessian determinant D_2 up to the $\operatorname{sign}[wR_2(p-1)R_1]$ (and $D_1 = 0$ identically); in the second pair of cases we have this sign is the sign of the Hessian determinant D_1 up to the $\operatorname{sign}[wR_2(p-1)R_1]$ (and $D_2 = 0$ identically).

By the way, we call the attention of the reader to the fact, that, for example, in 4_1) above we have necessarily $y_2 < 0$ (here y_2 is fixed and our extremal trajectories in the plane (y_1, y_3) here hit $y_1 = -y_2 = |y_2|$ as we are always under restrictions $-y_1 \le y_2 \le y_1$, that is $y_1 \ge |y_2|$), so $(-2y_2)^{p-3}$ makes a perfect sense. The same type of observation holds for all other cases.

To complete the investigation of $\operatorname{sign} D_i$ we need to calculate the sign of the expression in the brackets in (15.49):

$$\omega R_2 + (p-1)R_1 = R_1^2[(p-1)\Phi' - \omega\Phi''] \tag{15.50}$$

$$
3_1) \qquad \begin{aligned} \Phi(\omega) &= G(\omega,-1),\\ \Phi'(\omega) &= G_{z_1}(\omega,-1) = p(\omega-1)^{p-2}(\omega+p-2),\\ \Phi''(\omega) &= G_{z_1z_1}(\omega,-1) = p(p-1)(\omega-1)^{p-3}(\omega+p-3),\\ (p-1)\Phi'-\omega\Phi'' &= -p(p-1)(p-2)(\omega-1)^{p-3}; \end{aligned}
$$

$$
4_1) \qquad \begin{aligned} \Phi(\omega) &= G(1,-\omega),\\ \Phi'(\omega) &= -G_{z_2}(1,-\omega) = -p(p-1)\omega(1-\omega-1)^{p-2},\\ \Phi''(\omega) &= G_{z_2z_2}(1,-\omega) = p(p-1)(1-\omega)^{p-3}[1-(p-1)\omega],\\ (p-1)\Phi'-\omega\Phi'' &= -p(p-1)(p-2)\omega(1-\omega)^{p-3}; \end{aligned}
$$

$$
3_2) \qquad \begin{aligned} \Phi(\omega) &= G(\omega,1),\\ \Phi'(\omega) &= G_{z_1}(\omega,1) = p(\omega+1)^{p-2}(\omega-p+2),\\ \Phi''(\omega) &= G_{z_1z_1}(\omega,1) = p(p-1)(\omega+1)^{p-3}(\omega-p+3),\\ (p-1)\Phi'-\omega\Phi'' &= -p(p-1)(p-2)(\omega+1)^{p-3}; \end{aligned}
$$

$$
4_2) \qquad \begin{aligned} \Phi(\omega) &= G(1,\omega),\\ \Phi'(\omega) &= G_{z_1}(1,\omega) = -p(p-1)\omega(\omega+1)^{p-2},\\ \Phi''(\omega) &= G_{z_1z_1}(\omega,-1) = -p(p-1)(\omega-1)^{p-3}[1+(p-1)\omega],\\ (p-1)\Phi'-\omega\Phi'' &= -p(p-1)(p-2)\omega(\omega+1)^{p-3}\,. \end{aligned}
$$

By the way, we call the attention of the reader to the fact, that, for example, in 4_1) above we have necessarily $y_2<0$, so by (15.34) $\omega<1$, so $(1-\omega)^{p-3}$ is fine there. The same type of observation works for other cases above. We see that in all cases $\operatorname{sign}[(p-1)\Phi'-\omega\Phi'']=-\operatorname{sign}(p-2)$. Therefore in the first two cases we have $D_2<0$ and this solution satisfies neither requirement (15.9) nor requirement (15.10). In the second two cases we have $D_1>0$, and the function M can be a candidate either for $\mathcal{M}_{\max}$ or for $\mathcal{M}_{\min}$ depending on the sign of the second derivative $M_{y_3y_3}$.

Recall that (see (15.48))

$$
M_{y_3y_3}=\frac{p\omega^{p-2}R_1H^2}{y_3^3}[\omega R_2+(p-1)R_1]
$$

Since in the case 3_2) we have $\operatorname{sign}[\omega R_2+(p-1)R_1]=-\operatorname{sign}(p-2)$ we need only to know $\operatorname{sign}R_1=\operatorname{sign}\Phi'=\operatorname{sign}\frac{d}{d\omega}G(\omega,1)$ Since this solution is considered only in the sector $\frac{p-2}{p}y_1 < y_2 < y_1$ (see Fig. 1), we have

$$G(y_1+y_2, y_1-y_2) = (2y_1)^{p-1}[py_2-(p-2)y_1] > 0\,, \tag{15.51}$$

and ω, being the unique positive solution of the equation

$$G(\omega,1) = (\omega+1)^{p-1}[\omega-p+1] = \frac{1}{y_3}G(y_1+y_2,y_1-y_2)\,, \tag{15.52}$$

satisfies the condition $\omega > p-1$. Therefore, $\operatorname{sign} R_1 = \operatorname{sign}\frac{d}{d\omega}G(\omega,1) = \operatorname{sign} p(\omega+1)^{p-2}(\omega-p+2) > 0$, and so $\operatorname{sign} M_{y_3y_3} = -\operatorname{sign}(p-2)$, i.e., for $p>2$ this is candidate for $\mathcal{M}_{\max}$ and for $p<2$ this is candidate for $\mathcal{M}_{\min}$.

We are still considering the case 3_2). Recall that this function is defined not in the whole domain Ξ_+, but only in the sector $\frac{p-2}{p}y_1 < y_2 < y_1$. To get a solution everywhere we need to "glue" this solution with that we obtained considering the case 2) (see (15.26)):

$$M(y) = (y_1+y_2)^p + C(y_3-(y_1-y_2)^p)\,. \tag{15.53}$$

To glue this solution along the plane $y_2 = \frac{p-2}{p}y_1$ with that we just obtained, let us require from the resulting function to be continuous everywhere. From (15.52) and (15.51) we see that $G(\omega,1) = 0$ on this plane. Therefore, $\omega = p-1$ and $M = \omega^p y_3 = (p-1)^p y_3$. The same value has solution (15.53) on this plane for $C = (p-1)^p$.

Now we need to check that we get correct continuation in the sense that if the solution satisfies (15.9), then its continuation satisfies the same condition as well, if the solution satisfies (15.10), then the same is true for its continuation. The Hessian determinants will have the right sign automatically (actually $D_2 = 0$ identically). We need only to check the sign of

$$M_{y_1y_1} = M_{y_2y_2} = p(p-1)\big((y_1+y_2)^{p-2} - (p-1)^p(y_1-y_2)^{p-2}\big)$$

in the domain $-y_1 < y_2 < \frac{p-2}{p}y_1$, or in the initial coordinates $0 < x_2 < (p-1)x_1$.

For $p>2$ we have

$$\begin{aligned}(y_1+y_2)^{p-2} = x_2^{p-2} &< (p-1)^{p-2}x_1^{p-2}\\ &< (p-1)^p x_1^{p-2} = (p-1)^p(y_1-y_2)^{p-2}\,,\end{aligned}$$

and for $p<2$ we have

$$(y_1+y_2)^{p-2}=x_2^{p-2}>(p-1)^{p-2}x_1^{p-2}$$
$$>(p-1)^p x_1^{p-2}=(p-1)^p(y_1-y_2)^{p-2}\,.$$

This means that M is a candidate for $\mathcal{M}_{\max}$ if $p>2$ and a candidate for $\mathcal{M}_{\min}$ if $p<2$, as it has to be.

Let us rewrite expression (15.53) in the same form, as it was made in (15.52).

$$M-Cy_3=(y_1+y_2)^p-C(y_1-y_2)^p=x_2^p-Cx_1^p\,. \tag{15.54}$$

Therefore, if we change a bit the definition of G defining it on the quadrant $z_i\ge 0$ as follows

$$G_p(z_1,z_2)=\begin{cases} z_1^p-(p-1)^p z_2^p, & \text{if } z_1\le (p-1)z_2\,,\\ (z_1+z_2)^{p-1}\big[z_1-(p-1)z_2\big], & \text{if } z_1\ge (p-1)z_2\,, \end{cases} \tag{15.55}$$

then we can write two our solutions M on Ξ_+ in an implicit form as before:

$$G(y_1+y_2,y_1-y_2)=y_3G(\omega,1)\,.$$

or solutions B on Ω_+

$$G(x_2,x_1)=G(B^{\frac1p},x_3^{\frac1p})\,, \tag{15.56}$$

In the case 4_2) we again consider exactly the same G_p from (15.55). In a similar way we can glue continuously the solution in case 4_2) found in the sector $-y_1< y_2<\frac{2-p}{p}y_1$

$$G(1,\omega)=(\omega+1)^{p-1}[1-(p-1)\omega]=\frac{1}{y_3}G(y_1-y_2,y_1+y_2)\,, \tag{15.57}$$

which is the same as

$$G(x_1,x_2)=G(x_3^{\frac1p},B^{\frac1p})\,, \tag{15.58}$$

with the solution (15.53) along the line $y_2=\frac{2-p}{p}y_1$. Here we have to take $C=(p'-1)^p$, because on the line $y_2=\frac{2-p}{p}y_1$ we have $G(1,\omega)=0$, i.e., $\omega=p'-1$. Now, in the sector $-y_1<y_2<\frac{2-p}{p}y_1$ we have

$$M_{y_3y_3}=\frac{p\omega^{p-2}R_1^2H^2}{y_3^3}\cdot\frac{p-2}{\omega+1}\,.$$

Therefore, $\operatorname{sign} M_{y_3y_3}=\operatorname{sign}(p-2)$, i.e., for $p<2$ this is candidate for $\mathcal{M}_{\max}$ and for $p>2$ this is candidate for $\mathcal{M}_{\min}$.

In the "dual" sector $x_2 > (p'-1)x_1$ (or $y_2 > \frac{2-p}{p}y_1$) for $p>2$ we have

$$\begin{aligned}(y_1+y_2)^{p-2} = x_2^{p-2} &> (p'-1)^{p-2}x_1^{p-2}\\ &> (p'-1)^p x_1^{p-2} = (p'-1)^p(y_1-y_2)^{p-2}\,,\end{aligned}$$

and for $p<2$ we have

$$\begin{aligned}(y_1+y_2)^{p-2} = x_2^{p-2} &< (p'-1)^{p-2}x_1^{p-2}\\ &< (p'-1)^p x_1^{p-2} = (p'-1)^p(y_1-y_2)^{p-2}\,.\end{aligned}$$

This means that M is a candidate for $\mathcal{M}_{\max}$ if $p<2$ and a candidate for $\mathcal{M}_{\min}$ if $p>2$.

Using the same "generalized" definition (15.55) of the function G we can write our solutions M on Ξ_+ in an implicit form as before:

$$G(y_1-y_2, y_1+y_2) = y_3 G(1,\omega)\,.$$

or solutions B on Ω_+

$$G(x_1,x_2) = G(x_3^{\frac{1}{p}}, B^{\frac{1}{p}})\,, \tag{15.59}$$

which should give, as we said above, the candidate for $\mathbf{B}_{\max}$ for $p<2$ and $\mathbf{B}_{\min}$ for $p>2$. Notice that for $p>2$ the candidate for, say, $\mathbf{B}_{\max}$ is given by equation (15.56).

It is a bit inconvenient to use one equation for, say, $B_{\max}$ if $p>2$ (this will be (15.56)), and another one (this will be(15.59)) for the same $\mathbf{B}_{\max}$ if $p<2$. We note that after interchanging role of z_i and replacing p by p' we get the scalar multiple of the original expression in both lines of (15.55). This allows us to give one expression for $B_{\max}$ for all p using notation of $p^* = \max\{p,p'\}$. In such a way we come to formula (13.4) for F_p, where we introduce additional scalar coefficients to make this function not only continuous but C^1-smooth everywhere in Ω_+. This smoothness guarantee us that the solution B is C^1-smooth as well.

16. Proof of Theorem 32. Verification theorem.

Exactly in the spirit of Stochastic Optimal Control theory we wrote the PDE (15.11), we solved it in the previous section by building B which solves the equations of Theorem 32 (these are the same equations as (15.56), (15.58)). Now continuing in the spirit of general results of Stochastic Optimal Control theory [FR], [WF] we need to prove that these solutions in fact are equal to $\mathbf{B}_{\max}, \mathbf{B}_{\min}$. In Stochastic

Optimal Control theory such proofs are called *verification theorems*, and they state roughly that if the solutions have a certain smoothness (often even slightly less than C^2), and if the domain is convex, then we are fine.

From now on we denote by $B_{\max}$ the unique positive solution of the equation $F(|x_2|,|x_1|) = F(B^{\frac{1}{p}}, x_3^{\frac{1}{p}})$ and by $B_{\min}$ the unique positive solution of the equation $F(|x_1|,|x_2|) = F(x_3^{\frac{1}{p}}, B^{\frac{1}{p}})$, where the function $F = F_p$ is defined in (13.4). Existence and uniqueness of the solution follows from the fact that $F(z_1, z_2)$ is strictly increasing in z_1 from $-p^{*(p-1)}(p^*-1)^p z_2^p$ till $+\infty$ as z_1 runs from 0 to $+\infty$ and it is strictly decreasing in z_2 from $p(p^*-1)^{p-1}z_2^p$ till $-\infty$ as z_2 runs from 0 to $+\infty$. Indeed, the first partial derivatives of F are

$$F_{z_1} = \begin{cases} pz_1^{p-1}, & \text{if } z_1 \le (p^*-1)z_2\,, \\ p(1-\frac{1}{p^*})^{p-1}(z_1+z_2)^{p-2}\big[pz_1 - \big((p-1)(p^*-1)-1\big)z_2\big], & \\ & \text{if } z_1 \ge (p^*-1)z_2\,; \end{cases} \tag{16.1}$$

$$F_{z_2} = \begin{cases} -(p^*-1)^p p z_2^{p-1}, & \text{if } z_1 \le (p^*-1)z_2\,, \\ -p(1-\frac{1}{p^*})^{p-1}(z_1+z_2)^{p-2}\big[(p^*-p)z_1 + p(p^*-1)z_2\big], & \\ & \text{if } z_1 \ge (p^*-1)z_2\,. \end{cases} \tag{16.2}$$

Note that both derivatives are continuous everywhere (even at the origin, where they vanish). Moreover, $F_{z_1} > 0$ if $z_1 > 0$ and $F_{z_2} < 0$ if $z_2 > 0$, i.e., F is strictly increasing in z_1 and strictly decreasing in z_2.

In the case of $B_{\max}$ we look for a solution of the equation

$$F(B^{\frac{1}{p}}, x_3^{\frac{1}{p}}) = F(|x_2|,|x_1|)$$

or

$$F(\omega, 1) = \frac{1}{x_3}F(|x_2|,|x_1|)\,.$$

Thus, we get a continuous solution $\omega(x)$ everywhere except the plane $x_3 = 0$, where ω is not defined. But we can easily estimate the behavior of ω nearly the line $x_3 = x_1 = 0$. Since F is decreasing in z_2 and $0 \le |x_1| \le x_3^{\frac{1}{p}}$, we have

$$F\Big(\frac{|x_2|}{x_3^{1/p}}, 1\Big) \le F(\omega, 1) = F\Big(\frac{|x_2|}{x_3^{1/p}}, \frac{|x_1|}{x_3^{1/p}}\Big) \le F\Big(\frac{|x_2|}{x_3^{1/p}}, 0\Big)\,.$$

Since F is increasing in z_1, we get

$$\frac{|x_2|}{x_3^{1/p}} \le \omega \le \omega_0\,,$$

where ω_0 is the solution of the equation

$$(\omega_0+1)^{p-1}(\omega_0-p^*+1)=\frac{|x_2|^p}{x_3}\,.$$

Whence $\omega_0 \ge p^*-1$ and

$$(\omega_0-p^*+1)^p \le (\omega_0+1)^{p-1}(\omega_0-p^*+1)=\frac{|x_2|^p}{x_3}\,,$$

i.e.,

$$\omega_0 \le p^*-1+\frac{|x_2|}{x_3^{1/p}}\,,$$

Therefore, for $B=\omega^p x_3$ we have the following estimate

$$|x_2|^p \le B \le \big(|x_2|+(p^*-1)x_3^{1/p}\big)^p\,,$$

which gives the continuity near $x_3=0$. Thus, the solution $B_{\max}$ is continuous in the closed domain Ω.

Similar considerations gives us the continuity of $B_{\min}$. In that case we have the equation

$$F(1,\omega)=\frac{1}{x_3}F(|x_1|,|x_2|)\,,$$

and hence

$$F\big(0,\frac{|x_2|}{x_3^{1/p}}\big) \le F(1,\omega)=F\big(\frac{|x_1|}{x_3^{1/p}},\frac{|x_2|}{x_3^{1/p}}\big) \le F\big(1,\frac{|x_2|}{x_3^{1/p}}\big)\,.$$

Now $F(1,\omega)$ is decreasing in ω, therefore,

$$\frac{|x_2|}{x_3^{1/p}} \le \omega \le \omega_0\,,$$

where ω_0 is the solution of the equation

$$1-(p^*-1)^p\omega_0^p=-(p^*-1)^p\frac{|x_2|^p}{x_3}\,,$$

i.e., $\omega_0^p=(p^*-1)^{-p}+|x_2|^p/x_3$ and for $B=\omega^p x_3$ we have the following estimate

$$|x_2|^p \le B \le |x_2|^p+(p^*-1)^{-p}x_3\,,$$

which gives the continuity near $x_3=0$. Thus, the solution $B_{\min}$ is continuous in the closed domain Ω as well.

First step of the proof is to check that the the main inequality (concavity (15.7) for the candidate $B_{\max}$ and convexity (15.8) for the candidate $B_{\min}$) is fulfilled if the points x^+, x^- satisfy the extra condition on their coordinates:

$$|x_1^+ - x_1^-| = |x_2^+ - x_2^-|\,. \tag{16.3}$$

This was almost done in the preceding section, when constructing these candidates. We know that the Hessians of our candidates have the required signs everywhere in our convex domain Ω except, possibly, the planes $x_1 = 0$, $x_2 = 0$, and, either $|x_2| = (p^*-1)|x_1|$ for $B_{\max}$ or $|x_1| = (p^*-1)|x_2|$ for $B_{\min}$. On these hyperplanes our solutions are not C^2-smooth, but this does not prevent them from being correctly concave (for the 3_2), $p > 2$ and 4_2), $p < 2$ cases) and correctly concave for the rest of the cases (namely, for the 3_2), $p < 2$ and 4_2), $p > 2$ cases). This one checks just by calculating directly the sign of the jump of the derivative. Namely, one fixes the line $L_t = a + bt$ in the direction of the vector $b = (b_1, b_2, b_3)$ such that $|b_1| = |b_2|$. We need to prove the concavity of B, the candidate for $\mathbf{B}_{\max}$, and the convexity of B, the candidate for $\mathbf{B}_{\min}$ on L_t. At any point of L_t, which is *not* the intersection of L_t with the abovementioned hyperplanes, this concavity (convexity) follows from the previous section, this is how the candidates for $\mathbf{B}_{\max}$, $\mathbf{B}_{\min}$ were built in (15.56), (15.58). At the points of intersections of L_t with the hyperplanes one can check the sign of the jump of the derivative of $B(a + tb)$. We leave this as an exercise for the reader.

Now we have in the cases (15.56), $p > 2$, and (15.58), $p < 2$, the following main inequality

$$B(x) - \frac{B(x^-) + B(x^+)}{2} \geq 0, \forall x^-, x^+ :$$

$$|x_1^+ - x_1^-| = |x_2^+ - x_2^-|, x_3^+ \geq |x_1^+|^p, x_3^- \geq |x_1^-|^p, x = \frac{x^+ + x^-}{2}\,. \tag{16.4}$$

And we have in the cases (15.56), $p < 2$, and (15.58), $p > 2$, the following main inequality

$$B(x) - \frac{B(x^-) + B(x^+)}{2} \geq 0, \forall x^-, x^+ :$$

$$|x_1^+ - x_1^-| = |x_2^+ - x_2^-|, x_3^+ \geq |x_1^+|^p, x_3^- \geq |x_1^-|^p, x = \frac{x^+ + x^-}{2}\,. \tag{16.5}$$

Theorem 38. *If B satisfies the main inequality* (16.4) *then $B \geq \mathbf{B}_{\max}$. If it satisfies* (16.5)*, then $B \leq \mathbf{B}_{\min}$.*

Proof. Let $I=[0,1]$ and J denote the lattice of its dyadic subintervals. As always $J_=, J_-$ are two sons of J. Let f,g be any two bounded measurable test functions on I such that $|(g,h_J)|=|(f,h_J)|$ for any J. Put

$$x=(\langle f\rangle_J,\langle g\rangle_J,\langle |f|^p\rangle_J),$$

$$x^+=(\langle f\rangle_{J_+},\langle g\rangle_{J_+},\langle |f|^p\rangle_{J_+}),$$

$$x^-=(\langle f\rangle_{J_-},\langle g\rangle_{J_-},\langle |f|^p\rangle_{J_-})\,.$$

The fact that $|(g,h_J)|=|(f,h_J)|$ exactly guarantees that x^+,x^- satisfy the assumptions of (16.4). So use (16.4) with such x,x^+,x^- for $J=I$, then for $J=I_=, J=I_-$, et cetera... We continue doing that for all dyadic J's, $|J|\ge 2^{-m+1}$. The sons of these smallest J's will be called $J_i^m, i=1,..,2^m$.

Adding up all our inequalities we get

$$\frac{1}{2^m}\sum_{i=1}^{2^m}B(\langle f\rangle_{J_i^m},\langle g\rangle_{J_i^m},\langle |f|^p\rangle_{J_i^m})\le B(x)\,.$$

Notice that $|\langle f\rangle_{J_i^m}|^p\to\langle |f|^p\rangle_{J_i^m}$ when $m\to\infty$. We observed that then

$$B(\langle f\rangle_{J_i^m},\langle g\rangle_{J_i^m},\langle |f|^p\rangle_{J_i^m})\to|\langle g\rangle_{J_i^m}|^p\,.$$

In other words we can write the left hand side as $\int_0^1 F_m(t)\,dt$, where non-negative F_m's converge almost everywhere to $|g(t)|^p$. By Fatou's lemma we obtain

$$\langle |g|^p\rangle_I\le B(x)\,,$$

which means exactly that $\mathbf{B}_{\max}(x)\le B(x)$.

For the case of (16.5), and $\mathbf{B}_{\min}$ we can use Lebesgue dominant convergence theorem instead of Fatou's lemma. We will get $\mathbf{B}_{\min}(x)\ge B(x)$ for B satisfying (16.5).

□

We are left to prove the opposite inequality

$$\mathbf{B}_{\max}\ge B(x)$$

for B from (15.56), $p>2$, and (15.58), $p<2$. (And similarly for $\mathbf{B}_{\min}$ with obvious changes.) This one can do by reversing the reasoning in the theorem above using the fact that domain $\Omega=\{x=(x_1,x_2,x_3): x_3\ge |x_1|^p\}$ is foliated by the straight line segments (extremal trajectories). The reader can see how this type of reasoning is done in [**?**]. The main idea is to travel along the extremal trajectories starting from $x \in \Omega$ to build a net $\mathcal{N} := \{x^+, x^-, x^{++}, x^{+-}, x^{--}, x^{-+}, ..., x = \frac{x^+ + x^-}{2}, x^+ = \frac{x^{++} + x^{+-}}{2}, x^- = \frac{x^{-+} + x^{-+}}{2}, \dots$. All points of the net should belong to Ω, and we put them on the same extremal trajectory on which x lies for a while. If one of them, say, z hits the boundary: $\partial\Omega$ (parabola) we stop building children z^+, z^-. But then one of them, say, ζ can hit (or approach) the special hyperplanes $x_1 = 0$ or $x_2 = 0$. In this case we choose *arbitrary* ζ^+, ζ^- to have $\zeta = \frac{\zeta^+ \zeta^-}{2}$ but in such a way that they lie in different quadrants. We also choose them very close to ζ, say δ_ζ close. Then we start anew a building of the net for ζ^+ and ζ^- separately. The reader can address to [VaVo2] to understand how the net $\mathcal{N}$ generates a pair of functions (f, g) such that

$$\mathbf{B}_{\max}(x) \geq B(x) - \epsilon\,,$$

where ϵ can be chosen as small as we wish by the choice of small δ_ζ's above.

This is a rough description of the way to prove $\mathbf{B}_{\max} \geq B(x)$, but essentially it is a shadow of a certain general verification theorem.

17. Function u_p from function $\mathbf{B}$

We found Burkholder's functions $\mathbf{B}_{\max}$ and $\mathbf{B}_{\min}$ as claimed in Theorem 32. As a corollary we immediately we get the sharp constant in Burkholder's inequality:

Theorem 39. *Let $I = [0, 1]$, $\langle f \rangle_I = x_1, \langle g \rangle_I = x_2$, g is a Martingale transform of f, $|x_2| \leq |x_1|$. Then*

$$\langle |g|^p \rangle_I \leq (p^* - 1)^p \langle |f|^p \rangle_I\,.$$

The constant $p^* - 1$, where $p^* := \max(p, \frac{p}{p-1})$ is sharp.

Proof. We just analyze the form of function $\mathbf{B}_{\max}$ from Theorem 32 and immediately see that

$$\sup_{x \in \Omega,\, |x_2| \leq |x_1|} \frac{B(x_1, x_2, x_3)}{x_3} = (p^* - 1)^p\,.$$

□

Theorem 40. *Let $I = [0, 1]$, $\langle f \rangle_I = x_1, \langle g \rangle_I = x_2$, g is a Martingale transform of f, $|x_2| \geq |x_1|$. Then*

$$\langle |f|^p \rangle_I \leq (p^* - 1)^p \langle |g|^p \rangle_I\,.$$

The constant $p^* - 1$, where $p^* := \max(p, \frac{p}{p-1})$ is sharp.

Proof. We just analyze the form of function $\mathbf{B}_{\min}$ from Theorem 32 and immediately see that

$$\inf_{x\in\Omega,\,|x_2|\ge|x_1|}\frac{B(x_1,x_2,x_3)}{x_3}=(p^*-1)^{-p}\,.$$

□

Remarks. 1) The same analysis shows that $\langle|g|^p\rangle_I\le(p^*-1)^p\langle|f|^p\rangle_I$ if and only if $|x_2|\le(p^*-1)|x_1|$ in Theorem 39, and in Theorem 40 $\langle|f|^p\rangle_I\le(p^*-1)^p\langle|g|^p\rangle_I$ if and only if $|x_2|\ge(p^*-1)^{-1}|x_1|$.

Notations. Below we use $\beta_p:=(p^*-1)^p$. Put

$$\phi_{\max}(x_1,x_2):=\sup_{x_3:(x_1,x_2,x_3)\in\Omega}[\mathbf{B}_{\max}(x_1,x_2,x_3)-\beta_p x_3]\,,$$

$$\phi_{\min}(x_1,x_2):=\inf_{x_3:(x_1,x_2,x_3)\in\Omega}[\mathbf{B}_{\min}(x_1,x_2,x_3)-\beta_p^{-1}x_3]\,.$$

These functions are defined on the whole $\mathbb{R}^2$.

Definition. We call the function ϕ on $\mathbb{R}^2$ zigzag concave if

$$\phi(x)-\frac{\phi(x^-)+\phi(x^+)}{2}\ge 0,\forall x,x^-,x^+\in\mathbb{R}^2:|x_1^+-x_1^-|=|x_2^+-x_2^-|,x=\frac{x^++x^-}{2}\,. \tag{17.1}$$

And we call the function ϕ on $\mathbb{R}^2$ zigzag convex if

$$\phi(x)-\frac{\phi(x^-)+\phi(x^+)}{2}\le 0,\forall x,x^-,x^+\in\mathbb{R}^2:|x_1^+-x_1^-|=|x_2^+-x_2^-|,x=\frac{x^++x^-}{2}\,. \tag{17.2}$$

The next theorem gives an independent description of $\phi_{\max}$ and $\phi_{\min}$.

Theorem 41. *Function $\phi_{\max}$ is the least zigzag concave majorant of function $h_p(x):=|x_2|^p-\beta_p|x_1|^p$. Function $\phi_{\min}$ is the greatest zigzag convex minorant of function $h_p(x):=|x_2|^p-\beta_p^{-1}|x_1|^p$.*

Remark. Notice that this is slightly counterintuitive: $\mathbf{B}_{\max}(x)-\beta_p x_3$ is sort of "concave", and the supremum of concave functions is *not* usually concave.

Proof. Let $x^-,x_+\in\mathbb{R}^2$ are such that $|x_1^+-x_1^-|=|x_2^+-x_2^-|$. Let $x=\frac{x^++x^-}{2}$. It is obvious that $\phi_{\max}$ is zigzag concave. One verifies this just by definition. In fact, if for $x^-\in\mathbb{R}^2$ the supremum in the definition of $\phi_{\max}$ is almost attained for x_3^- and correspondingly if for $x^+\in\mathbb{R}^2$ the supremum in the definition of $\phi_{\max}$ is almost attained for a certain x_3^+, we can define $x_3=\frac{x_3^-+x_3^+}{2}$, and use (16.4) to write

$$\mathbf{B}(x_1,x_2,x_3)-\beta_p x_3-\frac{B(x_1^-,x_2^-,x_3^-)-\beta_p x_3^-+B(x_1^+,x_2^+,x_3^+)-\beta_p x_3^+}{2}\ge 0\,.$$

Using the facts that $\phi(x_1,x_2) \geq \mathbf{B}(x_1,x_2,x_3) - \beta_p x_3$, and $\phi(x^-) \approx B(x_1^-, x_2^-, x_3^-) - \beta_p x_3^-$, $\phi(x^+) \approx B(x_1^+, x_2^+, x_3^+) - \beta_p x_3^+$, we conclude (17.1). Inequality (17.2) is totally similar.

As sup is bigger than lim we conclude

$$\phi_{\max}(x_1,x_2) \geq \lim_{x_3 \to |x_1|^p} [\mathbf{B}_{\max}(x_1,x_2,x_3) - \beta_p x_3] = |x_2|^p - \beta_p |x_1|^p =: h_p(x)\,.$$

As inf is smaller than lim we get analogously

$$\phi_{\min}(x_1,x_2) \leq \lim_{x_3 \to |x_1|^p} [\mathbf{B}_{\min}(x_1,x_2,x_3) - \beta_p x_3] = |x_2|^p - \beta_p |x_1|^p =: h_p(x)\,.$$

This is because the boundary values of $\mathbf{B}_{\max}$ and $\mathbf{B}_{\min}$ are $|x_2|^p$.

We are left to see that $\phi_{\max}$ is the *least* such majorant (and a symmetric claim for $\phi_{\min}$). Let ψ be a zigzag concave function such that

$$\phi_{\max} \geq \psi \geq h_p\,. \tag{17.3}$$

Consider function $\Psi(x_1,x_2,x_3) := \psi(x_1,x_2) + \beta_p x_3$. It is immediate that Ψ satisfies (16.4). On the boundary of Ω we have $\Psi(x) \geq |x_2|^p$, this is just by the right hand side of (17.3). Then exactly as in Theorem 38 we get that

$$\Psi(x) \geq \mathbf{B}_{\max}(x)\,.$$

Then, obviously,

$$\psi(x) = \sup_{x_3 : (x,x_3) \in \Omega} [\Psi(x) - \beta_p x_3] \geq \sup_{x_3 : (x,x_3) \in \Omega} [\mathbf{B}_{\max}(x) - \beta_p x_3] = \phi_{\max}(x)\,.$$

So we proved that $\phi_{\max}$ is the least zigzag concave majorant of h_p. Symmetric consideration will bring us the fact that $\phi_{\min}$ is the largest zigzag convex minorant of h_p.

□

The reader should look now at function F_p from Theorem 32. It would be interesting to obtain the formulae for $\phi_{\max}$ and $\phi_{\min}$, especially using this F_p. It would be also interesting to understand the role of function

$$u_p(x_1,x_2) := p(1 - \frac{1}{p^*})^{p-1} (|x_1| + |x_2|)^{p-1} (|x_2| - (p^* - 1)|x_1|)\,, \tag{17.4}$$

mentioned in the introduction and used repeatedly by Burkholder. May be it is equal to $\phi_{\max}$? The answer is "no", but we can prove the following

Theorem 42.

$$\phi_{\max}(x_1,x_2) = F_p(x_2,x_1)\,. \tag{17.5}$$

Also

Theorem 43.

$$\phi_{\min}(x_1,x_2) = -(p^*-1)^{-p}F_p(x_1,x_2)\,. \tag{17.6}$$

It is sufficient to prove only Theorem 42.

Proof. Let us consider the case $p>2$. Functions $\phi_{\max}$ and F_p are both p-homogeneous and symmetric with respect the change of x_1 to $-x_1$ and x_2 to $-x_2$. So it is enough to look at them on $S := \{x_1 > 0, x_2 > 0, x_1 + x_2 = 1\}$. We know the formula for $\mathbf{B}_{\max}$ from Theorem 32. It follows from this formula, that on $S\cap\{x_2 \le (p-1)x_1\}$ function $\phi_{\max}$ can be easily written down (for such (x_1,x_2) function $\mathbf{B}$ is $x_2^p-(p-1)^p(x_3-x_1^p)$ by our construction in Theorem 32). In fact, we get on $S\cap\{x_2\le(p-1)x_1\}$ that $\phi_{\max}(x_1,x_2) = x_2^p-(p-1)^p x_1^p$. The last expression coincides with $F_p(x_2,x_1)$ on $S\cap\{x_2\le(p-1)x_1\}$. On the rest of S function F_p is linear. On the other hand we can calculate $(\phi_{\max})''_{x_1x_1} - (\phi_{\max})''_{x_2x_2}$ on the rest of S by using the fact that we know $\mathbf{B}_{\max}$ from Theorem 32.

In fact, fix $x\in S\cap\{x_2>(p-1)x_1\}$. Let x_3 almost ensures the supremum in the definition of $\phi_{\max}(x)$: $B(x,x_3)-\beta_p x_3 \ge \phi_{\max}(x)-\epsilon$. Point (x,x_3) by construction lives on an extremal trajectory of $\mathbf{B}_{\max}$, on this trajectory (which is not vertical) function $\mathbf{B}_{\max}$ is linear. Choose x^+,x^-,x_3^+,x_3^- in such a way, that (x^+,x_3^+) and $)x^-,x_3^-$ are on the same trajectory, they are not the same as (x,x_3) but they are very close to it. This is possible as trajectory is not vertical. Also we can guarantee that $x=\frac{x^++x^-}{2}, x_3=\frac{x_3^++x_3^-}{2}$. We know three things:

$$\mathbf{B}(x,x_3)-\beta_p x_3 \ge \phi_{\max}(x)-\epsilon\,,$$

$$\mathbf{B}(x^+,x_3^+)-\beta_p x_3^+ \le \phi_{\max}(x^+)\,,$$

$$\mathbf{B}(x^-,x_3^-)-\beta_p x_3^- \le \phi_{\max}(x^-)\,.$$

But also the linearity of $\mathbf{B}$ on this trajectory gives

$$\mathbf{B}(x,x_3)-\beta_p x_3 - \frac{1}{2}[\mathbf{B}(x^+,x_3^+)-\beta_p x_3^+ + \mathbf{B}(x^-,x_3^-)-\beta_p x_3^-] = 0\,.$$

Combining all this we get

$$\phi_{\max}(x) - \frac{1}{2}[\phi_{\max}(x^+)+\phi_{\max}(x^-) \le \epsilon\,.$$

We already saw in Theorem 41 that $\phi_{\max}$ is zigzag concave. The previous inequality (with $\epsilon \to 0$) gives that $\phi_{\max}$ is linear on $S^+ := S \cap \{x_2 > (p-1)x_1\}$. This exactly how $F_p(x_2, x_1)$ behaves on this interval. Both functions are obviously continuous. Function $F_p(x_2, x_1)$ is a concave smooth function majorazing h_p on S. This is immediate from its formula. Functions $\phi_{\max}$ and $F_p(x_2, x_1)$ are equal on $S \setminus S^+$. And they are both continuous on S and linear on S^+. If suddenly $\phi_{\max} > F_p(x_2, x_1)$ on S^+ then these knowledge shows that it is not concave on S. This is a contradiction. If suddenly $\phi_{\max} < F_p(x_2, x_1)$ on S^+ then the smoothness of $F_p(x_2, x_1)$ shows that $\phi_{\max}$ becomes below h_p near the point $x_2 = (p-1)x_1$ on S. But it is a majorant of h_p. We come again to contradiction. Therefore, we proved $\phi_{\max}(x_1, x_2) = F_p(x_2, x_1)$.

□

17.1. **Function u_p.** Burkholder often used function u_p from (17.4). To demystify it let us notice that it is also p-homogeneous and as such can be considered only on S. On this segment function u_p becomes linear. It is a majorant of h_p, and it is tangent to h_p exactly at point $x_2 = (p-1)x_1$ on S, where h_p vanishes. It is not the least zigzag concave function greater than h_p (of course not, $\phi_{\max}$ is such), but it is the least zigzag concave function larger than h_p and such that on segment S (and all segments obtained by dilation tS and symmetries $x \to -x_1, x_2 \to -x_2$) it is not only concave, but also linear.

This is already proved, and we leave the detailed reasoning to the reader. One more thing we want to mention is that we could have considered a slightly more general problem. Namely, instead of majoraizing function $h_p(x_1, x_2) = |x_2|^p - (p^* - 1)^p |x_1|^p$ we could have started with any function

$$H_c(x_1, x_2) := |x_2|^p - c\,|x_1|^p\,.$$

The reader can easily see that we have proved the following theorem (of course Burkholder already proved the most of it long ago).

Corollary 44. *The smallest c for which there exists a zigzag concave function Φ_c majoraizing H_c is equal to $(p^* - 1)^p$. For this c the least zigzag majorant is $F_p(x_1, x_2)$. The smallest c or which there exists a zigzag concave function Φ_c majoraizing H_c such that it is linear on the segment S, symmetric and p-homogeneous* *is equal to* $(p^*-1)^p$. *For this* c *the least zigzag majorant linear on* S, *symmetric and* p*-homogeneous is* $u_p(x_1,x_2)$.

Remark. Notice an interesting thing which we do not know how to explain. Given function $\mathbf{B}_{\max}$ from Theorem 32 we can easily diminish the number of variables and construct ϕ_{max}. But amazingly we can also find $\mathbf{B}_{\max}$ if only $\phi_{\max}$ is given. In fact, Theorem 42 gives the formula for $\phi_{\max}$ via F_p. Then F_p allows us to find $\mathbf{B}_{\max}$. If we now combine Theorems reft1 and Theorem 42 to conclude that

Corollary 45. *Given a point* $x\in\Omega$ *we can find* $\mathbf{B}_{\max}(x)$ *by solving equation:*

$$\phi_{\max}(x_1,x_2)=\phi_{\max}(x_3^{\frac1p},\mathbf{B}_{\max}^{\frac1p})\,.$$

Symmetric formula allows to find $\mathbf{B}_{\min}$ *if* $\phi_{\min}$ *is known:*

$$\phi_{\min}(x_2,x_1)=\phi_{\min}(\mathbf{B}_{\min}^{\frac1p},x_3^{\frac1p})\,.$$

18. Ahlfors-Beurling operator. Calculus of variations. Morrey's and Sverak's problems.

"Everything has been thought of before, the task is to think about it again" said Goethe. We want to take another look at Ahlfors-Beurling operator T, it is the operator that sends $\bar\partial f$ to ∂f for smooth functions f with compact support on the plane $\mathbb{C}$. Here

$$\partial f=\frac{\partial f}{\partial z}=\left(\frac{\partial f}{\partial x}-i\frac{\partial f}{\partial x}\right),\ \bar\partial f=\frac{\partial f}{\partial z}=\left(\frac{\partial f}{\partial x}-+i\frac{\partial f}{\partial x}\right).$$

We intentionally omitted $\frac12$, this will not bring complications.

This operator was much studied in the last 30 years. There are several reasons for that.

a) Operator T and its multidimensional analogs play an important part in the theory of quasiregular mappings.

b) Attempts to calculate (estimate) the norm of T are closely related to important conjectures in the Calculus of Variation: Morrey's conjecture of 1952 and Sverak's conjecture of 1992. Morrey's conjecture states that "rank one convex functions are not necessarily quasiconvex", so in essence it asks for a series of counterexamples, of rank one convex functions that are not quasiconvex. Sverak's conjecture asks about a *concrete* rank one convex function whether it is quasiconvex.

c) There is a deep connection of Ahlfors–Beurling operator to stochastic calculus and stochastic optimal control.

Saying all that let us state several very innocent looking problems.

19. Some problems from the Calculus of Variation

We mostly follow in this section the exposition of A. Baernstein–S. Montgomery-Smith [BaMS].

Define a function $L : \mathbb{C}^2 \to \mathbb{R}$ as follows

$$L(z,w) = \begin{cases} |z|^2 - |w|^2\,, & \text{if } |z| + |w| \le 1\,, \\ 2|z| - 1\,, & \text{if } |z| + |w| > 1\,. \end{cases}$$

Sverak's problem: Let $f \in C_0^\infty(\mathbb{C})$. Is it true that

$$\int_{\mathbb{C}} L(\bar\partial f, \partial f)dxdy \ge 0\,? \tag{19.1}$$

We can restate this problem in the language of quasiconvex functions of matrix argument. Then we will explain Morrey's problem.

Let $M(m,n)$ be the set of all $m \times n$ matrices with real entries. A function $\Psi : M(m,n) \to \mathbb{R}$ is called *rank one convex* if $t \to \Psi(A + tB)$ is a convex function for any $B \in M(m,n)$ that has rank 1 and any $A \in M(m,n)$.

Function Ψ is called *quasiconvex* if it is locally integrable, and for each $A \in M(m,n)$ and each bounded domain $\Omega \subset \mathbb{R}^n$ and each smooth function $f : \mathbb{R}^n \to \mathbb{R}^m$ one has

$$\frac{1}{|\Omega|}\int_\Omega \Psi(A + Df)dx \ge \Psi(A)\,. \tag{19.2}$$

Here Df is the Jacobi matrix of the map f.

For $n = 1$ or $m = 1$ quasiconvexity is equivalent to convexity (which of course is equivalent for this case to rank one convexity). Always convexity implies quasiconvexity that implies rank one convexity.

Morrey's problem: If $m > 1, n > 1$ rank one convexity does not imply quasiconvexity. This was conjectured by Morrey in 1952 in [Mo]. Sverak [Sv2] proved that problem if $m > 2$. If $m = 2$ this is still open even in the case $n = 2$. Morrey's problem enjoyed a lot of attention in the last 57 years.

We can translate easily Sverak's problem to this language (this is how it appeared in the first place).

In fact, for $A\in M(2,2)$, $A=\begin{bmatrix} a & b\\ c & d\end{bmatrix}$. Put $z=a-d+i(b+c)$, $w=a+d+i(c-b)$. We see that

$$\Psi(A):=L(z,w)=\begin{cases}-4\det(A)\,, & \text{if } (|A|_2^2-2\det A)^{\frac12}+(|A|_2^2+2\det A)^{\frac12}\le 1\,,\\ 2(|A|_2^2-2\det A)^{\frac12}-1\,, & \text{otherwise}\,.\end{cases} \tag{19.3}$$

This function is rank one convex on $M(2,2)$. A very simple proof is borrowed from [BaMS]. We fix $A,B\in M(2,2)\,,\mathrm{rank}(B)=1$. Let (z,w) cooresponds to A and (Z,W) to B. The fact that $\mathrm{rank}(B)=1$ means that the map $\zeta\to Z\zeta+W\bar\zeta$ maps the plane to the line, so $|Z|=|W|$. Then $|z+tZ|^2-|w+tW|^2=a+tb$ for some $a,b\in\mathbb{R}$—there is no quadratic term. Also $\Psi(A+tB)=|z+tZ|^2-|w+tW|^2=a+tb$ if and only if $|z+tZ|+|w+tW|\le 1$. As all z,Z,w,W is fixed and $t\to|\alpha+t\beta|$ is convex for any complex α,β, we conclude that $\{t\in\mathbb{R}:|z+tZ|+|w+tW|\le 1\}$ is an interval (may be empty). On the other hand outside of this interval $\Psi(A+tB)=2|z+tZ|-1$, that is a convex function. Now continuity of $\Psi(A+tB)$ implies that it is convex.

Let $f:\mathbb{R}^2\to\mathbb{R}^2$, smooth, and with compact support. Write $f=u+iv$, $Df=\begin{bmatrix} u_x & u_y\\ v_x & v_y\end{bmatrix}$. Then using the notations above $z=u_x-v_y+i(v_x+u_y)=\bar\partial f$, $w=u_x+v_y+i(v_x-u_y)=\partial f$.

This means that $\Psi(Df)=L(\bar\partial f,\partial f)$ and Sverak's conjecture states that

$$\int\Psi(Df)dxdy\ge\Psi(0)=0\,.$$

In other words, (19.2) means that Ψ from (19.3) is quasiconvex at $A=0$. We conclude that would Sverak's conjecture not be true, then Ψ would give an example of rank one convex function which is not quasiconvex. This would solve Morrey's conjecture which exactly asks for such an example for the case $n=2,m=2$.

However, (19.2) is probably true. Everybody who worked with these questions believes in it. We will explain this belief.

20. Consequences of Sverak's inequality (19.2).

In what follows

$$p^*=\max(p,\frac{p}{p-1})=\max(p,p')\,.$$

Here is one other function on $M(2,2)$ which is rank one convex but for which it is unknown whether it is quasiconvex. It is also on $M(2,2)$. Several such functions are discussed in [Sv1], [Sv2], but the function Ψ above and Ψ_p below are especially important for us.

$$\Psi_p(A) =$$
$$((p^*-1)(|A|_2^2-2\det A)^{\frac{1}{2}}-(|A|_2^2+2\det A)^{\frac{1}{2}})((|A|_2^2-2\det A)^{\frac{1}{2}}+(|A|_2^2+2\det A)^{\frac{1}{2}})^{p-1}\,.$$

Repeat our correspondence between real matrices $M(2,2)$ and $\mathbb{C}^2$: for $A \in M(2,2)$, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, put $z = a-d+i(b+c)$, $w = a+d+i(c-b)$. We see that

$$\Psi_p(A) := L_p(z,w) = ((p^*-1)|z|-|w|)(|z|+|w|)^{p-1}\,. \tag{20.1}$$

See now e. g. [BaMS] for the following formula

$$L_p(z,w) = \frac{2}{p(2-p)}\int_0^\infty t^{p-1}L(\frac{z}{t},\frac{w}{t})dt\,,\ \text{ if } 1<p<2\,.$$

Obviously, for any $z, w, Z, W, |Z| = |W|$ the function $t \to L_p(z+tZ, w+tW)$ is convex because of the formula and because we just proved such a property for L.

Then, automatically,

$$\Psi_p(A) = \int_0^\infty t^{p-1}\Psi(\frac{A}{t})dt\,,\ \text{ if } 1<p<2\,.$$

And then Ψ_p is a rank one convex function in an obvious way, if $1<p<2$.

But for $2<p<\infty$ another formula holds (see again [BaMS]): put

$$M(z,w) = L(z,w)-(|z|^2-|w|^2) = (|w|^2-(|z|-1)^2)\mathbf{1}_{|z|+|w|>1}\,.$$

Obviously, for any $z, w, Z, W, |Z| = |W|$ the function $t \to M(z+tZ, w+tW)$ is convex because we substract the linear term $a+bt$ from $L(z+tZ, w+tW)$.

Let us define

$$\mathcal{M}_p(z,w) := \frac{2}{p(p-1)(p-2)}\int_0^\infty t^{p-1}M(\frac{z}{t},\frac{w}{t})dt\,,\ \text{ if } 2<p<\infty\,,$$

then it is such that for any $z, w, Z, W, |Z| = |W|$ the function $t \to M_p(z+tZ, w+tW)$ is convex.

And, automatically,

$$\Psi_p(A) = \int_0^\infty t^{p-1}(\Psi(\frac{A}{t})+\frac{4}{t^2}\det A)dt\,,\ \text{ if } 2<p<\infty\,,$$

is a rank one convex function on $M(2,2)$.

Banuelos-Wang problem: Is it true that for any smooth function with compact support on $\mathbb{C}$

$$\int_{\mathbb{C}} L_p(\bar{\partial} f, \partial f)dxdy \geq 0\,? \tag{20.2}$$

If (20.2) were *not* true we would have that Ψ_p is *not* quasiconvex at $A = 0$ and Morrey's problem would be solved in the remaining case.

If (20.2) were true than we would have solved Iwaniec's problem of 1982.

Iwaniec's problem: Ahlfors-Beurling operator T which sends $\bar{\partial} f$ to ∂f has norm $p^* - 1$. Essentially it is the following inequality for all $f \in C_0^\infty(\mathbb{C})$:

$$\int_C |\partial f|^p dxdy \leq (p^* - 1)^p \int_C |\bar{\partial} f|^p dxdy\,? \tag{20.3}$$

In equavalent form (20.3) is stated as follows

$$\int_C |Tf|^p dxdy \leq (p^* - 1)^p \int_C |f|^p dxdy\,, \text{ for all } f \in C_0^\infty(\mathbb{C})\,? \tag{20.4}$$

In fact, (20.2) $\Rightarrow$ (20.3) follows from a pioneering research of Burkholder, who in [Bu1], [Bu3], p. 77, noticed that

$$p\left(1 - \frac{1}{p^*}\right)^{p-1} L_p(z, w) \leq (p^* - 1)^p |z|^p - |w|^p\,. \tag{20.5}$$

Now it is clear why (20.2) implies (20.3).

Remark. What is subtle and interesting is the whole theory of inequalities of the type like Burkholder's inequlity (20.5). What really is going on in (20.5) is the search of function (in this case the left hand side), which posesses ceratin convexity properties, and at the same time posesses the "obstacle" property expressed by inequlity (20.5). We are looking for the 'somewhat convex" envelope of the right hand side. This is actually the essence of the so-called Bellman function approach. The literature is now extensive, and it relates (20.5) to Monge-Ampère equation and stochastic control, see e. g. Slavin-Stokols' paper [SlSt] or Vasyunin and Volberg [VaVo2].

Sverak's conjecture (19.1) and, as a result, Banuelos-Wang's conjecture (20.2) were proved in the paper of Baernstein and Montgomery-Smith [BaMS] in the case of so-called "stretch functions" f. A stretch function (in our notations, which differ slightly from those in [BaMS]) is a function of the form

$$f(re^{i\theta}) = g(r)e^{-i\theta}\,,$$

where g is a smooth function on $\mathbb{R}_+$, $g(0) = g(\infty) = 0$, and $g \geq 0$. We will call such g's **stretches**.

Therefore,

$$\int_C |Tf|^p dxdy \leq (p^* - 1)^p \int_C |f|^p dxdy\,, \text{ for all } f \in C_0^\infty(\mathbb{C})\,, f(z) = f(|z|)\,, f : \mathbb{C} \to \mathbb{R} \tag{20.6}$$

holds for all real valued radial f. This was a question in [BaJa2] whether this inequality holds for complex valued radial f.

We show how to do that using Bellamn function techniques. The advantage of this method is that it is applicable to other situations. It also illustrate how genuinely convex functions can sometimes be involved in a rather sophisticated way in proving *quasiconvexity* statements.

21. Bellman function and Ahlfors-Beurling operator on radial functions.

The kernel of T is $K(z) = \frac{1}{\pi}\frac{1}{z^2} =: e^{-2i\theta}k(r)\,, z = re^{i\theta}$. So for radial g

$$Tg(\rho e^{i\varphi}) = \int_{\mathbb{C}} K(z) g(|z - \rho e^{i\varphi}|) dA(z) = e^{-2i\varphi} \int\int e^{-2i\psi} k(r) g(|re^{i\psi} - \rho|)\, r dr d\psi =$$

$$e^{-2i\varphi} \int K(w + \rho) g(|w|) dA(w) = e^{-2i\varphi} \int_0^\infty (\int_0^{2\pi} K(|w|e^{it} + \rho)\, dt) g(|w|) |w| dt d|w|\,.$$

21.1. **Symmetrization.** If we denote $n(\rho, r) = \int_0^{2\pi} K(re^{it} + \rho)\, dt$, and $Ng(\rho) := \int_0^\infty n(\rho, r) g(r)\, r dr$ we get

$$Tg(\rho e^{i\varphi}) = e^{-2i\varphi} \int_0^\infty n(\rho, r) g(r)\, r dr =: e^{-2i\varphi} Ng(\rho)\,.$$

Hence to check the norm of $Tg(\rho e^{i\varphi})$ in L^p we can take a function $f \in L^{p'}(\mathbb{C})$, write *bilinear* form

$$(f, Tg) = \int_{\mathbb{C}} f Tg\, dA = \int (\int e^{-2i\psi} f(re^{i\psi}) d\psi) Ng(r)\, r dr\,.$$

Let us notice that the family $\mathcal{F}$ of functions having the form

$$f(re^{i\theta}) = \sum_{k=-N}^{N} e^{-ik\theta} f_k(r)\,,$$

where f_k are smooth compactly supported functions, give us a dense family in $L^{p'}(\mathbb{C})$, , $1 < p < \infty$. Also let us call $e^{-ik\theta} f_k(r)$ a k-mode function. The set of k-modes is called $\mathcal{F}_k$. Continuing the last formula we write

$$(f, Tg) = \int_{\mathbb{C}} f Tg\, dA =$$

$$2\pi \int f_{-2}(r) Ng(r)\, r dr = \int_{\mathbb{C}} f_{-2}(|z|) Ng(|z|)\, dA(z) = (e^{2i\theta} f_{-2}(|z|), Tg)\,. \qquad (21.1)$$

Let us notice that projection $\Pi_k : \mathcal{F} \to \mathcal{F}_k$ has norm at most 1 in any L^p. In fact, let R_φ is a rotation of $\mathbb{C}$ by φ. Then

$$\Pi_k f = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\varphi} f(R_\varphi z)\, d\varphi\,.$$

So projestion Π_k is just the averaging-type operator, and thus has norm 1.

Conclusion: to estimate $\|Tg\|_p$, $g \in \mathcal{F}_0$, it is sufficient to estimate the bilinear form $|(f, Tg)|$ only for $f \in \mathcal{F}_{-2}$ (and in the unit ball of $L^{p'}(\mathbb{C})$). We proved actually the following

Lemma 46. *For $g \in \mathcal{F}_k$,*

$$\|Tg\|_p = \sup_{f \in L^{p'} \cap \mathcal{F}_{k-2}\,, \|f\|_{p'} \le 1} |(f, Tg)|\,.$$

We actually repeated also the following well-known simple calculation.

Lemma 47. *Let a complex valued kernel $K(re^{i\theta}) = e^{-il\theta} k(r)$. Let $\mathcal{K} f := K \star f$ be a convolution operator. Then it maps $\mathcal{F}_k$ to $\mathcal{F}_{k-l}$ and for every $g = e^{-ik\theta} g_k(r) \in \mathcal{F}_k$ we have*

$$\mathcal{K} g(\rho\, e^{i\varphi}) = e^{-i(k-l)\varphi} \int_0^\infty N_k(\rho, r) g_k(r)\, r dr\,,$$

where

$$N_k(\rho, r) := \int_0^{2\pi} K(re^{it} + \rho) e^{-ikt}\, dt\,.$$

For $\mathcal{K} = T$ one can compute the kernel of N_m:

$$\frac{1}{2} N_m(t, x) = x\delta_x - (m+1) \frac{1}{t^{m+2}} x^m \mathbf{1}_{[0,t]}(x)\,.$$

It is not very nice, but let us denote by $h : \mathbb{R}_+ \to \mathbb{R}_+$ the map $h(t) = t^2$. Let us define the operator Λ_m (see [BaJa2] for this)

$$\Lambda_m g(u) = g(u) - (m+1) \frac{1}{u^{\frac{m+2}{2}}} \int_0^u v^{\frac{m}{2}}\, g(v)\, dv\,.$$

For $m = 0$ this is $\Lambda_0 = \mathrm{Id} - H$, where H is Hardy's averaging operator on half-axis:

$$Hg(u) := \frac{1}{u}\int_0^u g(v)\,dv\,.$$

Famous Hardy's inequality is practically equivalent to computing

$$\|H\|_{l^p(\mathbb{R}_+)\to L^p(\mathbb{R}_+)} = p^*\,,\ \text{if } 1 < p \le 2\,.$$

Curiously, we can see now that the question about complex valued radial functions from [BaJa2] is equivalent to

$$\|H - \mathrm{Id}\|_{l^p(\mathbb{R}_+)\to L^p(\mathbb{R}_+)} \le p^* - 1\,,\ \text{if } 1 < p \le 2\,. \tag{21.2}$$

21.2. **A Bellman function.** We will use a certain interesting convex functions on $\mathbb{R}^6$ and $\mathbb{R}^4$ to approach our "quasiconvexity" inequality (20.3) for complex valued radial functions.

Suppose we have function $\mathcal{B}(u, v, \xi, \eta, H, Z)$ of 6 real variables defined in

$$\Omega = \{|(u,v)|^p \le H\,, |(\xi,\eta)|^{p'} < Z\}\,,$$

and satisfying two conditions:

I) For an arbitrary $a \in \Omega\alpha \in \mathbb{R}^6$ we want to have

$$\langle -\frac{d^2\mathcal{B}}{da^2}\alpha, \alpha\rangle \ge 2(\alpha_1^2 + \alpha_2^2)^{1/2}(\alpha_3^2 + \alpha_4^2)^{1/2}\,.$$

and

II) Everywhere in Ω

$$0 \le \mathcal{B}(a) \le (p^* - 1)\left(\frac{H}{p} + \frac{Z}{p'}\right),\ \text{where } p^* = \max(p, p')\,,$$

For the sake of future convenience we prefer to work with the following transformation of $\mathcal{B}$ ($a = (u, v, \xi, \eta, H, Z)$):

$$B(u, v, \xi, \eta) := \sup_{H,Z:(u,v,\xi,\eta,H,Z)\in\Omega}\{\mathcal{B}(a) - (p^* - 1)\left(\frac{H}{p} + \frac{Z}{p'}\right)\}\,.$$

Then it is not difficult to check that this B is still concave (inspite of being *the supremum* of concave functions): In fact,

$$-d^2B \ge 2|(du, dv)||(d\xi, d\eta)|\,, \tag{21.3}$$

$$-(p^* - 1)\left(\frac{|(u,v)|^p}{p} + \frac{|(\xi,\eta)|^{p'}}{p'}\right) \le B(u, v, \xi, \eta) \le 0\,. \tag{21.4}$$

The existence of such $\mathcal{B}$ was proved in [PV],[DV1].

21.3. **Heat extension.** Let f, g be two test functions on the plane. By the same letters we denote their heat extensions into $\mathbb{R}^3_+$. This is a simple lemma observed in [PV]:

Lemma 48.

$$\int_{\mathbb{C}} fTg\, dA = -2\int_{\mathbb{R}^3_+} (\partial_x + i\partial_y)f \cdot (\partial_x + i\partial_y)g\, dxdydt\,.$$

Let us use below the following notations:

$$f = u + iv, z_1 = u_x + iu_y, z_2 = v_x + iv_y,$$

$$g = \xi + i\eta, \zeta_1 = \xi_x + i\xi_y, \zeta_2 = \eta_x + i\eta_y\,.$$

Now we can read Lemma 48 as follows:

$$\int fTg = -2\int_{\mathbb{R}^3_+} (z_1 + iz_2)(\zeta_1 + i\zeta_2)\,, |\int fTg| \le 2\int_{\mathbb{R}^3_+} |z_1 + iz_2||\zeta_1 + i\zeta_2|\,. \quad (21.5)$$

And from here we see

$$|\int fTg| \le 2\int \left[\frac{|z_1 + iz_2|^2 + |z_1 - iz_2|^2}{2}\right]^{1/2} \left[\frac{|\zeta_1 + i\zeta_2|^2 + |\zeta_1 - i\zeta_2|^2}{2}\right]^{1/2}. \quad (21.6)$$

Property (21.3) of B can be rewritten

Lemma 49.

$$-\langle d^2B(z_1, z_2, \zeta_1, \zeta_2)^T, (z_1, z_2, \zeta_1, \zeta_2)^T\rangle \ge 2[|z_1|^2 + |z_2|^2]^{1/2}[|\zeta_1|^2 + |\zeta_2|^2]^{1/2}\,.$$

This lemma gives now

$$-2\langle d^2B(\frac{z_1 + iz_2}{2}, \frac{z_1 - iz_2}{2}, \frac{\zeta_1 + i\zeta_2}{2}, \frac{\zeta_1 - i\zeta_2}{2})^T, (\text{the same vector})^T\rangle \ge$$
$$2\left[\frac{|z_1 + iz_2|^2 + |z_1 - iz_2|^2}{2}\right]^{1/2} \left[\frac{|\zeta_1 + i\zeta_2|^2 + |\zeta_1 - i\zeta_2|^2}{2}\right]^{1/2}.$$

After integration and using (21.6) we get

$$|\int_{\mathbb{C}} fTg| \le \int_{\mathbb{R}^3_+} LHS\,. \quad (21.7)$$

The rest is the estimate of $\int_{\mathbb{R}^3_+} LHS$ from above. First of all simple algebra ($a := (u, v, \xi, \eta)$:

$$\int_{\mathbb{R}^3_+} LHS = -\frac{1}{2}\int_{\mathbb{R}^3_+} \langle d^2B(a)(z_1, z_2, \zeta_1, \zeta_2)^T, (\text{the same})^T\rangle -$$
$$\frac{1}{2}\int_{\mathbb{R}^3_+} \langle d^2B(a)(z_2, -z_1, \zeta_2, -\zeta_1)^T, (\text{the same})^T\rangle +$$
$$\int_{\mathbb{R}^3_+} \text{auxilliary terms} =: I + II + III\,.$$

It has been proved in [PV], [DV1] that (the convention is that u, v, ξ, η are heat extensions of homonome functions on the plane)

$$I = \frac{1}{2}\int_{\mathbb{R}^3_+} \left(\frac{\partial}{\partial t} - \Delta\right) B(u, v, \xi, \eta)\,,$$
$$II = \frac{1}{2}\int_{\mathbb{R}^3_+} \left(\frac{\partial}{\partial t} - \Delta\right) B(v, -u, \eta, -xi)\,.$$

An estimate of I from above. Let H denote the heat extension of function $|f|^{p'} = (u^2 + v^2)^{p'/2}$, Z denote the heat extension of function $|g|^p = (\xi^2 + \eta^2)^{p/2}$. In $\mathbb{R}^3_+$ consider $\Psi(x, y, t) = B(u, v, \xi, \eta) + (p^* - 1)\left(\frac{H}{p'} + \frac{Z}{p}\right)$. Then

$$2I = \int_{\mathbb{R}^3_+} \left(\frac{\partial}{\partial t} - \Delta\right)\Psi\,,$$

Then obviously (integration by parts)

$$2I = \int_{\mathbb{R}^3_+} \frac{\partial}{\partial t}\Psi = \lim_{t\to\infty}\int_{\mathbb{R}^2} \Psi(\cdot, t) - \int_{\mathbb{R}^2} \Psi(\cdot, 0) \le \lim_{t\to\infty}\int_{\mathbb{R}^2} \Psi(\cdot, t)\,.$$

Using (21.4) ($B \le 0$) we get

$$I \le \frac{1}{2}(p^* - 1)\lim_{t\to\infty}\int_{\mathbb{R}^2} \left(\frac{H(\cdot, t}{p'} + \frac{Z(\cdot, t)}{p}\right) = \frac{1}{2}(p^* - 1)\left(\frac{\|f\|_{p'}^{p'}}{p'} + \frac{\|g\|_p^p}{p}\right).$$

Similarly,

$$II \le \frac{1}{2}(p^* - 1)\left(\frac{\|-if\|_{p'}^{p'}}{p'} + \frac{\|-ig\|_p^p}{p}\right).$$

So

$$I + II \le (p^* - 1)\left(\frac{\|f\|_{p'}^{p'}}{p'} + \frac{\|g\|_p^p}{p}\right).$$

We are going to prove next that

$$III \le 0\,.$$

Combining we get $|\int_{\mathbb{C}} fTg| \le (p^*-1)\Big(\frac{\|f\|_{p'}^{p'}}{p'} + \frac{\|g\|_p^p}{p}\Big)$ and the usual polarization argument proves out final statement:

$$|\int_{\mathbb{C}} fTg| \le (p^*-1)\|f\|_{p'}\|g\|_p\,.$$

21.4. **Why** $III = \int_{\mathbb{R}^3_+}$ **auxilliary terms** $dxdydt \le 0$**?** First of all the symmetry implies that

$$B(u,v,\xi,\eta) = \Phi(\sqrt{u^2+v^2}, \sqrt{\xi^2+\eta^2})\,.$$

So far we did not use the fact that

$$g(z) = \xi(r) + i\eta(r)\,, f(z) = e^{2i\theta}(m(r)+ik(r))\,. \tag{21.8}$$

Let as before $a(x,y,t) = (u,v,\xi,\eta)$ with heat extension functions. Automatically, with afixed t

$$\Phi(a), d\Phi(a), d^2\Phi(a)\,,\ \text{depend only on}\ \ r+\sqrt{x^2+y^2}\,. \tag{21.9}$$

Remark. In proving that $III = 0$ we are going to use this fact a lot. But $III = 0$ seems to hold under some other assumptions on f, g.

All auxiliary terms are in

$$\langle d^2B(a)(z_1,z_2,\zeta_1,\zeta_2)^T, (z_2,-z_1,\zeta_2,-\zeta_1)^T\rangle - \langle d^2B(a)(z_2,-z_1,\zeta_2,-\zeta_1)^T, (z_1,z_2,\zeta_1,\zeta_2)^T\rangle\,.$$

This expression $= A + D_1 + D_2 + C$, where

$$A = (B_{11}+B_{22})\Im z_2\bar z_1\,, C = (B_{33}+B_{44})\Im\zeta_2\bar\zeta_1\,.$$

Also

$$D_1 = B_{13}\Im\zeta_2\bar z_1 + B_{23}\Im\zeta_2\bar z_2 + B_{14}\Im z_1\bar\zeta_1 + B_{24}\Im z_2\bar\zeta_1\,.$$

$$D_2 = B_{13}\Im z_2\bar\zeta_1 + B_{23}\Im\zeta_1\bar z_1 + B_{14}\Im z_2\bar\zeta_2 + B_{24}\Im\zeta_2\bar z_1\,.$$

Why $\int_{\mathbb{R}^2} D_1(x,y,t)\,dxdy = 0$**?**

In D_1 the smaller index of B_{kl}, $k \in 1,2, l \in 3,4$ coincides with the index of z_i. In D_2 this is not the case. This is the explanation why integrating each term of D_1 returns 0. For exmple, (the last equality uses $\eta_\theta = 0$)

$$\Im\zeta_2\bar{z}_1 = \det\begin{bmatrix} u_x, \, \eta_x \\ u_y, \, \eta_y \end{bmatrix} = \det\begin{bmatrix} u_r, \, \eta_r \\ u_\theta/r, \, \eta_\theta/r \end{bmatrix} = -\eta_r u_\theta/r\,.$$

But (recall $f = u + iv, g = \xi + i\eta$)

$$B_{13} = \frac{u}{|f|}\frac{\xi}{|g|}\Phi_{12}(|f|,|g|)\,.$$

Then the first term of D_1

$$= \phi(r)uu_\theta\,,$$

and its integral along any circle is zero. Similarly,

$$D_1 = (uu_\theta + vv_\theta)\frac{\eta\xi_r - \xi\eta_r}{r}\frac{\Phi(|f|,|g|)}{|f||g|}\,, \tag{21.10}$$

and so $D_1 = \phi_1(r)uu_\theta + \phi_2(r)vv_\theta$. Hence for each fixed t

$$\int_{\mathbb{R}^2} D_1(x,y,t)\,dxdy = 0\,.$$

Coming to D_2 we can simalarly see that

$$D_2 = (uv_\theta - vu_\theta)\frac{\xi\xi_r + \eta\eta_r}{r}\frac{\Phi_{12}(|f|,|g|}{|f||g|}\,. \tag{21.11}$$

Recall that from (21.8) it follows that $u = m\cos 2\theta - k\sin 2\theta\,, v = m\sin 2\theta + k\sin 2\theta$, and from this

$$uv_{theta} - vu_\theta = 2(m^2(r) + k^2(r)) = 2(u^2(r) + v^2(r)) = 2|f|^2(r) =: 2M^2(r)\,.$$

Using similarly the notation $N(r) = |g|$ we can see from (21.11) and the previous equality that

$$D_2 = \frac{2}{r}\Phi_{12}(M(r),N(r))N'(r)M(r)\,.$$

Now we compute

$$A = (B_{11} + B_{22})\Im z_2\bar{z}_1 = (\Phi_{11} + \frac{1}{M}\Phi_1)(u_r v_\theta/r - v_r u_\theta/r)\,.$$

Using (21.8) we get $u_r v_\theta - v_r u_\theta = 2M(r)M'(r)$. Therefore

$$A = \frac{2}{r}(\Phi_{11} + \frac{1}{M}\Phi_1)MM' = \frac{2}{r}(\Phi_{11}MM' + \Phi_1 M')\,.$$

Notice (again (21.8)) that in

$$C = (B_{33} + B_{44})\Im\zeta_2\bar{\zeta}_1$$

the expression $\Im\zeta_2\bar\zeta_1 = \xi_r\eta_\theta/r - \eta_r\xi_\theta/r = 0$. So $C = 0$.

Adding the expressions for A, D_2 we obtain after integration over $\mathbb{R}^2$:

$$\int_{\mathbb{R}^2}(A(x,y,t)+D_2(x,y,t))\,dxdy = 4\pi\int_0^\infty(\Phi_12(M,N)M(r)N'(r)+$$
$$\Phi_{11}(M,N)M(r)M'(r))\,dr + 4\pi\int_0^\infty \Phi_1(M,N)M'(r)\,dr =: a+b\,.$$

Integrating b by parts we get $-a$ and $-\Phi_1(M(0),N(0))M(0)$. Function Φ grows in the first varible. So

$$\int_{\mathbb{R}^3_+}\text{auxilliary terms}\,dxdydt \le 0\,,$$

and we are done.

Actually, in our particular case $III = 0$. Function $f = u+iv$ on $\mathbb{C}$ has the form $f = e^{2i\theta}(m(r)+ik(r))$, therefore, its heat extension $f(x,y,t)$ obviously satisfies $f(x,y,0) = 0$. So $M(0) = 0$. As we saw

$$III = -\Phi_1(M(0),N(0))M(0) = 0\,.$$

References

[B1] J. Bourgain *Some remarks on Banach spaces in which martingale difference sequences are unconditional*, Ark. Mat. **21**(1983), 163–168.

[BaJa1] R. Banuelos, P. Janakiraman, *L^p–bounds for the Beurling–Ahlfors transform.* Trans. Amer. Math. Soc. **360** (2008), no. 7, 3603–3612.

[BaJa2] R. Banuelos, P. Janakiraman,

[BaMH] R. Bauelos, P. J. Mndez-Hernndez, *Space-time Brownian motion and the Beurling-Ahlfors transform.* Indiana Univ. Math. J. **52** (2003), no. 4, 981–990.

[BaWa1] R. Banuelos, G. Wang, Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms, Duke Math. J., **80** (1995), 575–600.

[BaMS] A. Baerstein and S. Montgomry-Smith, *Some conjectures about integral means of ∂f and $\overline{\partial} f$*, Complex anlysis and Differential Equations, Proc. of the Marcus Wallenberg symposium in honour of Matts Essén, Uppsala, Sweeden, 1997, 92–109.

[BBCH] J. Bennett, N. Bez, A. Carbery, D. Hundertmark *Heat flow and Strichartz.* Personal communication.

[BCT] J. Bennett, A. Carbery, T. Tao *On the multilinear restriction and Kakeya conjectures,* arxiv: math/0509262v1 12 Sep 2005.

[BV] M. Bateman, A. Volberg, *The power law for the Buffon needle probability of the four-corner Cantor set: an estimate from below*, Preprint, 2008, pp. 1–10.

[Bu1] D. BURKHOLDER, *Boundary value problems and sharp estimates for the martingale transforms*, Ann. of Prob. **12** (1984), 647–702.

[Bu2] D. BURKHOLDER, *An extension of classical martingale inequality*, Probability Theory and Harmonic Analysis, ed. by J.-A. Chao and W. A. Woyczynski, Marcel Dekker, 1986.

[Bu3] D. BURKHOLDER, *Sharp inequalities for martingales and stochastic integrals*, Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987), Astérisque No. 157-158 (1988), 75–94.

[Bu4] D. BURKHOLDER, *Differential subordination of harmonic functions and martingales*, (El Escorial 1987) Lecure Notes in Math., **1384** (1989), 1–23.

[Bu5] D. BURKHOLDER, *Explorations of martingale theory and its applications*, Lecture Notes in Math. **1464** (1991), 1–66.

[Bu6] D. BURKHOLDER, *Strong differential subordination and stochastic integration*, Ann. of Prob. **22** (1994), 995–1025.

[Bu7] D. BURKHOLDER, *A proof of the Peczynski's conjecture for the Haar system*, Studia MAth., **91** (1988), 79–83.

[Dac] B. DACOROGNA, *Some recent results on polyconvex, quasiconvex and rank one convex functions*, Adv. Math. Appl. Sci., World Science Publ., 1994, pp. 169-176.

[Dac1] B. DACOROGNA, Direct Methods in the Calculus of Variations, Springer, 1989.

[DV1] O. DRAGICEVIC, A. VOLBERG *Sharp estimates of the Ahlfors-Beurling operator via averaging of Martingale transform,* Michigan Math. J. **51** (2003), 415-435.

[DV2] O. DRAGICEVIC, A. VOLBERG, *Bellman function, Littlewood–Paley estimates, and asymptotics of the Ahlfors–Beurling operator in $L^p(\mathbb{C})$, $p \to \infty$*, Indiana Univ. Math. J. **54** (2005), no. 4, 971–995.

[DV3] O. DRAGICEVIC, A. VOLBERG, *Bellman function and dimensionless estimates of classical and Ornstein-Uhlenbeck Riesz transforms.* J. of Oper. Theory, **56** (2006) No. 1, pp. 167-198.

[DPV] O. DRAGICEVIC, S. PETERMICHL, A. VOLBERG,*A rotation method which gives linear L^p-estimates for powers of the Ahlfors-Beurling operator.* Journal des Mathématiques Pures et Appliqueés, **86**, No. 6 (2006), 492-509.

[DTV] O. DRAGICEVIC, S. TREIL, A. VOLBERG, *A lemma about* 3 *quadratic forms*, arXiv:0710.3249. To appear in Intern. Math. Research Notices.

[FR] W. FLEMING, R. RISHEL, Deterministic and Stochastic Optimal Control. Applications of Mathematics. Stochastic Modelling and Applied Probability Vol. 1, Springer Verlag, 1998.

[WF] W. FLEMING, *Controlled Markov processes and viscosity solutions of nonlinear evolution*, Publications of the Scuola Normale Superiore, 2007.

[GMSS] S. GEISS, S. MONTGOMERY-SMITH, E. SAKSMAN, On singular integral and martingale transforms, arxiv: math. CA/0701516v1 18 Jane 2007.

[Gar] J. B. GARNETT, "Bounded analytic functions," Acad. Press, NY, 1981.

[G-CRdF] J. Garcia-Cuerva, J. Rubio de Francia, Weighted Norm Inequalities and Related Topics, North-Holland, 1985.

[GPTV1] A. Gillespie, S. Pott, S. Treil and A. Volberg, *Logarithmic growth for matrix martingale transform*, Journal of London Math. Soc., **64** (2001), 624–636.

[GV] R. F. Gundy, N. Th. Varopoulos, , *Les transformations de Riesz et les intgrales stochastiques*, C. R. Acad. Sci. Paris Sr. A-B 289 (1979), no. 1, A13–A16.

[HJ] V. P. Havin and B. Jöricke , *The Uncertainity Principle in Harmonic Analysis,* Springer-Verlag, Berlin, 1994.

[Hu] S. Hukovic, Thesis, Brown University, 1998.

[HTV] S. Hukovic, S. Treil, A. Volberg, *Bellman functions and sharp weighted estimates for the square functions.* Complex analysis, operators, and related topics (in memory of S.A.Vinogradov), Oper. Theory Adv. Appl., **113**, Birkhauser, 2000, pp. 97–113.

[Ib] I. A. Ibragimov, Completely regular multidimensional stationary processes with discrete time, *Proc. Steklov Inst. Math.*, **111** (1970), 269–301.

[KV] N. Kalton, I. Verbitsky

[M] A. S. Markus, Introduction to the Spectral Theory of Polynomial Operator Pencils, American Mathematical Society, 1988.

[Mo] Morrey

[My] P. A. Meyer, *Transformations de Riesz pour les lois Gaussiennes*, Lect. Notes. Math., **1059** (1984), 79–193.

[N1] F. Nazarov, *A counterexample to Sarason's conjecture*, Preprint, MSU, 1997, 19 pp; Available at `http://www.math.msu.edu/~fedja/prepr.html`

[N2] F. Nazarov, *Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type* St. Petersburg Math. J. **5** (1994), no. 4, 663–717.

[NPTV] F. Nazarov, G. Pisier, S. Treil and A. Volberg , *Sharp estimates in vector Carleson imbedding theorem and for vector papraproducts*, J. Reine Angew. Math. 542 (2002), 147–171.

[NT] F. Nazarov, S. Treil, *The hunt for Bellman function: applications to estimates of singular integral operators and to other classical problems in harmonic analysis, Algebra i Analysis* **8** (1997), no. 5, 32-162.

[NTV3] F. Nazarov, S. Treil, A. Volberg, *A counterexample to infinitely dimensional Carleson embedding theorem*, C.R. Acad.Sci., Paris, **325** (1997), no. 4, 383-388.

[NTV4] F. Nazarov, S. Treil, A. Volberg, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. of Amer. Math. Soc., 12 (1999), 909-928.

[NTV8] F. Nazarov, S. Treil, A. Volberg,*Bellman function in stochastic control and harmonic analysis*, Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000), 393–423, Oper. Theory Adv. Appl., 129, Birkhauser, Basel, 2001.

[NTV10] F. Nazarov, S. Treil, A. Volberg,*Two weight estimate for the individual Haar multipliers and and other well localized operators.* Preprint, 2004, pp. 1–14.

[NV] F. NAZAROV AND A. VOLBERG, *Bellman function, two weighted Hilbert transforms and embeddings of the model spaces K_θ*, Dedicated to the memory of Thomas H. Wolff. J. Anal. Math. 87 (2002), 385–414.

[NV1] F. NAZAROV AND A. VOLBERG, *Heating of the Ahlfors–Beurling operator and estimates of its norm*, St. Petersburg Math. J., **14** (2003) no. 3.

[Ni] N. K. NIKOLSKI, *Treatise on the shift operator,* Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences],**273** (1986), Springer-Verlag, Berlin-New York.

[NiT] N. NIKOLSKI AND S. TREIL, *Linear resolvent growth of rank one perturbation of a unitary operator does not imply its similarity to a normal operator*, Dedicated to the memory of Thomas H. Wolff. J. Anal. Math. 87 (2002), 415–431.

[NiB] N. E. BENAMARA, N. K. NIKOLSKI, *Resolvent tests for similarity to a normal operator*, Proc. London. Math. Soc., **78**(1999), no.3, pp. 585–626.

[P] S. PETERMICHL*A sharp bound for weighted Hilbert transform in terms of classical A_p charactreristic*, Amer. J. Math. 129 (2007), no. 5, 1355–1375.

[Pi] S. PICHORIDES, *On the best values of the constants in the theorems of M. Riesz, Zygmund, and Kolmogorov*, Studia Math., **44** (1972) 165–179.

[PTV] S. PETERMICHL, S. TREIL, A. VOLBERG, *Why the Riesz transforms are averages of the dyadic shifts?* Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000). Publ. Mat. 2002, Vol. Extra, 209–228.

[PV] S. PETERMICHEL, A. VOLBERG, *Heating the Beurling operator: weakly quasiregular maps on the plane are quasiregular,* Duke Math. J., **112** (2002), no.2, pp. 281–305.

[Pog] A. V. Pogorelov, *Extrinsic geometry of convex surfaces.* Translations of Mathematical Monographs, Amer. Math. Soc., v. 35, 1973.

[Saw1] E. SAWYER, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math., **75** (1982), no. 1, pp. 1–11.

[Saw2] E. SAWYER, *Two–weight norm inequalities for certain maximal and integral operators*, Lecture Notes Math., **908**, Springer-Verlag, Berlin–Heidelberg–New York, 1982, pp. 102–127.

[SlSt] L. Slavin, A. Stokolos, *The maxial operator on $L^p(\mathbb{R}^n)$*, Preprint, 2007.

[SV] L. Slavin, V. Vasyunin, *Sharp results in the integral-form John–Nirenberg inequality*, Preprint, 2007; http://arxiv.org/abs/0709.4332.

[St] E. STEIN, Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, with the assistance of Timothy S. Murphy, Princeton Math. Ser. **43**, Monographs in Harmonic analysis, iii, Princeton Univ. Press, Princeton, 1993.

[Str] D. STROOK, *Probability Theory, an Analytic View*, Cambridge University Press, 1993.

[StInt] I. KARATZAS, S. SHREVE, Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics, Springer, 1991.

[Sv1] V. SVERAK, *Rank-one convexity does not imply quasiconvexity*, Proc. Royal Soc. Edinburg, v. 120A, (1990), 185-189.

[Sv2] V. Sverak, *New examples of quasiconvex functions*, Arch. Rational Mech. Anal., v. 119, (1992), 293-300.

[TV1] S. Treil and A. Volberg, *Wavelets and the angle between Past and Future*, J. of Funct. Analysis, **143**, (1997), no. 2, pp. 269–308.

[TV2] S. Treil and A. Volberg, *Completely regular multivariate stationary processes and the Muckenhoupt condition*, Pacific J. Math., **190**, (1999), no. 2, 361–382.

[TVZ] S. Treil, A. Volberg, D. Zheng, *Hilbert transform, Toeplitz operators and Hankel oerators, and invariant A_∞ weights*, Revista Mat. Iberoamericana, **13**, (1997), no. 2, 319–360.

[Varo1] N. Varopoulos,

[Va] V. Vasyunin, *The sharp constant in the John–Nirenberg inequality*, Preprint PDMI no.20, 2003; http://www.pdmi.ras.ru/preprint/2003/index.html.

[Ve] J. Verdera, *L^2 boundedness of the Cauchy integral and Menger curvature*, Cont. Math., **277**, Amer. Math. Soc., Providence, RI, 2001, pp. 139–158.

[Vo1] A. Volberg, *Matrix A_p weights via S-functions*, J. Amer. Math. Soc., **10**, (1997), no. 2, pp. 445–466.

[Vo2] A. Volberg, Calderón-Zygmund Capacities and Operators on Non-homogeneous Spaces, CBMS Lecture Notes in Math., **100**, 2003, pp. 1-167, Amer. Math. Soc, Providence, RI.

[VoEcole] A. Volberg, *Bellman approach to some problems in Harmonic Analysis*, Séminaires des Equations aux derivées partielles. Ecole Politéchnique, 2002, exposé XX, pp. 1–14.

[VaVo] V. Vasyunin and A. Volberg, *Monge–Ampère equation and Bellman opimization of Carleson embedding theorems*, Amer. Math. Soc. Transl. (2), vol. 226, "Linear and Complex Analysis", 2009, pp. 195–238.

[Vo] A. Volberg, *Bellman approach to some problems in Harmonic Analysis*, Seminaire des Equations aux derivées partielles, Ecole Polytéchnique, 2002, exposé XX, 1–14.

[VaVo1] V. Vasyunin, A. Volberg, *The Bellmanfunction for certain two weight inequality: the case study*, St. Petersburg Math. J. **18** (2007), no. 2, pp. 201–222.

[VaVo2] V. Vasyunin, A. Volberg, *Monge–Ampère equation and Bellman optimization of Carleson Embedding Theorems*, arXiv:0803.2247.

[VaVo3] V. Vasyunin, A. Volberg, *Bellster and others*, Preprint, 2008.

[Ve] *A. D. Ventzel, Introduction to the theory of stochastic processes.*

[W] J. Wittwer *A sharp estimate on the norm of the martingale transform.* Math. Res. Lett. **7** (2000), no. 1, 1–12.

Vassili Vasyunin, V.A. Steklov. Inst.
vasyunin@pdmi.ras.ru

Alexander Volberg, Department of Mathematics, Michigan State University and the University of Edinburgh. volberg@math.msu.edu and a.volberg@ed.ac.uk